

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Validation of discharge diagnosis codes to identify serious infections among older adults

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020857
Article Type:	Research
Date Submitted by the Author:	27-Nov-2017
Complete List of Authors:	Wiese, Andrew; Vanderbilt University Medical Center, Health Policy Griffin, Marie R; Vanderbilt University Medical Center, Health Policy Stein, Michael; Vanderbilt University, Pharmacology Schaffner, William; Vanderbilt University Medical Center, Health Policy Greevy, Robert; Vanderbilt School of Medicine, Biostatistics Mitchel, Jr., Edward; Vanderbilt University Medical Center, Health Policy Grijalva, Carlos; Vanderbilt University, Health Policy
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Infectious diseases, Epidemiology
Keywords:	coding algorithms, Medicaid, older adults, serious infections

SCHOLARONE™
Manuscripts

1
2
3 1 **Validation of discharge diagnosis codes to identify serious infections among older adults**
4

5 2

6
7 3 **Running title:** Validation of diagnosis codes to identify infections
8

9 4

10
11 5 **Authors:** Andrew D. Wiese, PhD, MPH¹; Marie R. Griffin^{1,2}, MD, MPH; C. Michael Stein, MB, ChB³;
12 6 William Schaffner, MD¹; Robert Greevy, PhD⁴; Edward F. Mitchel Jr., MS¹; Carlos G. Grijalva, MD,
13 7 MPH^{1,2}

14
15
16
17 8 **Affiliations:** ¹Department of Health Policy, Vanderbilt University School of Medicine, Nashville,
18 9 Tennessee, USA; ²Mid-South Geriatric Research Education and Clinical Center, VA Tennessee Valley
19 10 Health Care System, Nashville, Tennessee, USA; ³Departments of Pharmacology and ⁴Biostatistics,
20 11 Vanderbilt University School of Medicine, Nashville, Tennessee, USA

21
22 12 **Corresponding Author:** Andrew D. Wiese, PhD, MPH; Department of Health Policy, Vanderbilt
23 13 University Medical Center, Suite 2600, Village at Vanderbilt, 1500 21st Avenue South, Nashville, TN
24 14 37212; phone: (615) 875-7997; email: andrew.d.wiese@vanderbilt.edu
25 15

26
27 16 **Key words:** coding algorithms; Medicaid; older adults; serious infections
28 17

29
30 18 **Word Count:** 3,411 /4,000
31 19

32
33 19 **Tables and Figures:** (3/5)
34 20
35 21
36 22
37 23
38 24
39 25
40 26

1
2
3 27 **ABSTRACT (262/300)**
4

5 28 **Objectives:** Hospitalizations for serious infections are common among older adults and frequently used
6
7 29 as study outcomes. Yet few studies have evaluated the performance of diagnosis codes to identify serious
8
9 30 infections in this population. We sought to determine the positive predictive value (PPV) of diagnosis
10
11 31 codes for identifying hospitalizations due to serious infections among older adults.

12
13 32 **Setting and participants:** We identified hospitalizations for possible infection among adults ≥ 50 years
14
15 33 enrolled in the Tennessee Medicaid healthcare program (2008-2013) using ICD-9 diagnosis codes for
16
17 34 pneumonia, meningitis/encephalitis, bacteremia/sepsis, cellulitis/soft-tissue infections, endocarditis,
18
19 35 pyelonephritis and septic arthritis/osteomyelitis.

20
21
22 36 **Design:** Medical records were systematically obtained from hospitals randomly selected from a stratified
23
24 37 sampling framework based on geographical region and hospital discharge volume.

25
26 38 **Measures:** Two trained clinical reviewers used a standardized extraction form to abstract information
27
28 39 from medical records. Pre-defined algorithms served as reference to adjudicate confirmed infection-
29
30 40 specific hospitalizations. We calculated the PPV of diagnosis codes using confirmed hospitalizations as
31
32 41 reference. Sensitivity analyses determined the PPV robustness to definitions that required radiological or
33
34 42 microbiological confirmation. We also determined interrater reliability between reviewers.

35
36
37 43 **Results:** The PPV of diagnosis codes for hospitalizations for infection (n=716) was 90% (95% CI: 88-
38
39 44 92). The PPV was highest for pneumonia [97% (95% CI: 95-98)] and cellulitis [91% (95% CI: 86-96)],
40
41 45 and lowest for meningitis/encephalitis [50% (95% CI: 19-81)]. The adjudication reliability was excellent
42
43 46 [93% agreement; first agreement-coefficient: 0.91]. The overall PPV was lower when requiring
44
45 47 microbiological confirmation [45%] and when requiring radiological confirmation for pneumonia [79%].

46
47 48 **Conclusions:** Discharge diagnosis codes have a high PPV for identifying hospitalizations for serious
48
49 49 infections among older adults, especially for common infections.
50

51 50

52 51

53 52

53 STRENGTHS AND LIMITATIONS OF THE STUDY

- 54 • This study examined the performance of diagnosis coding algorithms to identify hospitalizations
55 due to serious infections among older adults enrolled in a State Medicaid program using a
56 systematic and representative sample of records from hospitals of different sizes and in distinct
57 State regions.
- 58 • The reference criteria to identify true infections was based on previous literature and clinical
59 expertise but may be imperfect. Nevertheless, identifying microbiologically-confirmed infections
60 is difficult due to the low sensitivity of culture-based diagnostic methods often used in clinical
61 practice.
- 62 • Diagnosis codes were based on the ICD-9-coding system only. These findings will continue to be
63 helpful for retrospective studies that encompass periods of ICD-9 use, yet additional studies
64 evaluating the performance of ICD-10-based codes would be beneficial.
- 65 • Our coding algorithms to identify serious infections had a high positive predictive value overall,
66 and will be useful in ongoing and future research using administrative data

78 INTRODUCTION

79 Infectious diseases remain a leading cause of morbidity and mortality in the U.S. and elsewhere
80 (1). Older adults, in particular, are at high risk for serious infections and their long-term consequences (2,
81 3). Among older adults, community-acquired serious infections (including pneumonia, sepsis, and
82 meningitis) often require hospitalization and represent a substantial burden on the U.S. healthcare system
83 (4-7). Therefore, it is important to monitor the incidence of these infections, identify important risk
84 factors, and determine the impact of preventative policies (e.g., vaccination) on these diseases among
85 older adults (8-10).

86 Large-scale epidemiological studies using administrative data often use serious infections as
87 outcomes (11-15). However, few studies have evaluated the performance of diagnosis codes to identify
88 serious infections among older adults. Most previous studies that have assessed the performance of coded
89 discharge diagnosis codes to identify serious infections have focused mainly on common infections (e.g.,
90 pneumonia or sepsis), specific populations (e.g., patients with rheumatoid arthritis), or on healthcare-
91 associated or hospital-acquired infections (16-25). Nevertheless, the performance of coded discharge
92 diagnoses for accurately identifying infections requiring hospitalization among older adults is unclear.
93 Therefore, we sought to determine the positive predictive value (PPV) of specific discharge diagnoses for
94 identifying infections that required hospitalization among older adults.

96 METHODS

97 Data sources

98 TennCare is the managed Medicaid program in the State of Tennessee that provides healthcare
99 insurance to those who are Medicaid eligible (around 20% of the Tennessee population). The adult
100 TennCare population consists of low-income pregnant women and individuals who are elderly or have a
101 disability (over 600,000 annually) (26). We used data from TennCare, supplemented with data from the
102 Tennessee Hospital Discharge System (a registry for all hospitalizations in Tennessee) and pharmacy
103 information from Medicare Part D for those that were dual eligible, to identify a retrospective cohort of

1
2
3 104 TennCare enrollees ≥ 50 years of age with pharmacy benefits (2008-2013). Cohort members had at least
4
5 105 180 days of baseline continuous enrollment before cohort entry, and were also required to be free of
6
7 106 certain life-threatening conditions known to increase the risk of infection (solid organ transplantation,
8
9 107 end-stage renal disease, HIV/AIDS, malignancy and serious kidney, liver and respiratory disease), have
10
11 108 evidence of at least one pharmacy prescription fill and evidence of at least one healthcare encounter
12
13 109 during baseline (to ensure detection of healthcare usage). Follow-up started on the earliest date the
14
15 110 inclusion criteria were met and continued through the earliest of the following: study end date (December
16
17 111 31, 2013), the day prior to diagnosis of a serious life-threatening condition that would have precluded
18
19 112 entry to the cohort, loss of enrollment, or date of death. From this retrospective cohort, we identified
20
21 113 possible hospitalizations for serious infections (see *Identification of hospitalizations for serious infection*)
22
23 114 for our validation study. To avoid including infections that may have originated due to a previous hospital
24
25 115 stay, we excluded hospitalizations for infections that occurred in the 30-day period after discharge from a
26
27 116 previous hospitalization.

117 **Identification of hospitalizations for serious infection**

118 Clinical knowledge and a literature review were used to identify primary discharge diagnosis
119 codes that have been used previously to identify specific serious infections that require hospitalization
120 (*study infections*), including pneumonia (alone or with a primary diagnosis of bacteremia/sepsis),
121 bacteremia/sepsis, pyelonephritis, meningitis/encephalitis, osteomyelitis/septic arthritis, endocarditis and
122 cellulitis (25, 27-29). Specific International Classifications of Diseases-Clinical Modification 9th-revision
123 (ICD-9-CM) diagnosis codes used to identify possible hospitalizations for each infection type are
124 presented in Table 1.

125 **Sampling Strategy**

126 We used stratified random sampling to select a representative subset of study infection
127 hospitalizations from among all possible cases identified in the retrospective cohort. Since larger hospitals
128 would be over represented in a purely random sampling, and because there may also be regional
129 variability in coding practices and infection prevalence, we constructed a sampling framework where

1
2
3 130 hospitals were stratified based on their geographic region in Tennessee (West, Central, and East), and
4
5 131 tertiles of reported discharge volume (Low, Medium, and High) during the study period (30-32). From
6
7 132 this sampling framework, we randomly selected three hospitals from each of these nine sampling strata,
8
9 133 and retrieved their medical records for review and validation (Figure 1). This strategy, relative to a purely
10
11 134 random sample, ensured better representation of infections identified in smaller hospitals and those in
12
13 135 more rural regions of the State of Tennessee. If a hospital refused to participate, it was replaced by
14
15 136 another hospital randomly selected from the same sampling stratum.

17
18 137 The overall goal was to review and validate 675 hospitalizations for serious infection from 27
19
20 138 hospitals (25 hospitalizations for each of the 3 hospitals comprising a stratum, yielding 75 hospitalizations
21
22 139 for each of the 9 strata) (Figure 1). We conservatively assumed that up to 80% of records requested
23
24 140 would be available for review, and so we requested 32 records per hospital to receive an average of 25
25
26 141 records from each (Figure 1). To ensure that we reviewed sufficient rare infections, we preferentially
27
28 142 selected any identified possible hospitalizations for meningitis/encephalitis, osteomyelitis/septic arthritis
29
30 143 and endocarditis from each hospital in the sample. We randomly selected the remaining set of possible
31
32 144 hospitalizations for other serious infections based on the proportional distribution of common infections
33
34 145 at each hospital (pneumonia, bacteremia/sepsis, pyelonephritis and cellulitis) until 32 infections were
35
36 146 identified. For hospitals with fewer than 32 infections during the study period, all infections were
37
38 147 requested.

41 148 **Abstraction of Medical Records**

42
43 149 Relevant clinical information was abstracted from the medical record (transfer notes, emergency
44
45 150 room summary, admission summary, physical/history, pharmacy information, laboratory, microbiology,
46
47 151 and radiology information, and discharge summary) of each hospitalization with a primary discharge
48
49 152 diagnosis code indicative of infection using a standardized and customized REDCap electronic data
50
51 153 capture instrument hosted at Vanderbilt University (33). As we were interested in infections that led to
52
53 154 hospitalizations, we focused our reviews on clinical, microbiological and radiological information from
54
55 155 the 2 days prior to the admission date through 2 days after admission to limit the possibility of identifying

1
2
3 156 infections that developed during the hospitalization (i.e. nosocomial infections). In preparation for this
4
5 157 study, the case report form was pilot-tested among a separate, convenience sample of 354 possible
6
7 158 infections identified in the cohort from 3 hospitals in the same city as Vanderbilt University. This separate
8
9 159 sample of hospitalizations was used only for pilot-testing the case report form, and was not included in
10
11 160 the current study. One trained medical reviewer abstracted the relevant information for all selected
12
13 161 records using the case report form. During the abstraction process, the lack of a particular finding in the
14
15 162 medical record was treated as a lack of evidence for that finding, and so no information was considered
16
17 163 missing after abstraction.

19 20 164 **Adjudication of Medical Records**

21
22 165 All records received were abstracted, reviewed and adjudicated. We made the final determination
23
24 166 of whether a hospitalization represented a confirmed infection or not using *a priori* definitions of clinical,
25
26 167 radiological, and/or microbiological findings compatible with infection for each infection type. Previous
27
28 168 validation studies and expert clinical knowledge were used to define these specific *a priori* definitions for
29
30 169 each infection type (*Supplementary appendix*) (25, 28, 34).

31 32 33 170 **Statistical analysis**

34
35 171 We calculated the PPV of the ICD-9-CM discharge diagnosis codes for identifying
36
37 172 hospitalizations for serious infection using the results of the *a priori* definitions applied to the information
38
39 173 abstracted from the medical records as the reference. Secondary analyses assessed the PPV for
40
41 174 hospitalizations for serious infection across hospitals of different sizes and in different geographical
42
43 175 regions of Tennessee.

44
45 176 We also assessed the reliability of the abstraction process. A second trained medical reviewer
46
47 177 abstracted relevant information from a subset of selected records, which included all meningitis and
48
49 178 endocarditis records, and a random selection of 10% of each of the remaining infection types. Each
50
51 179 reviewer conducted the process independently and blinded from one another. For the subset of records
52
53 180 abstracted by both reviewers, inter-reviewer agreement for the adjudication of a true or mis-identified
54
55 181 hospitalization was assessed using the Gwet's first agreement coefficient (AC_1) (36-38). Since Cohen's

1
2
3 182 kappa statistic can be unreliable when the prevalence of the event and the level of observer agreement are
4
5 183 high in the study sample, we used Gwet's AC₁ as a reliability measure unlikely to be affected by these
6
7 184 concerns (38-40). In sensitivity analyses, we assessed the impact of excluding hospitalizations that
8
9 185 occurred after the individual was transferred from another healthcare facility, as initial documentation and
10
11 186 details of the infection could be missing or incomplete in the receiving hospital (34). We also assessed the
12
13 187 impact on the PPV for all infections when requiring microbiological identification of a pathogen
14
15 188 (excluding common contaminants) from a sterile site within 2 days before or after the hospitalization
16
17 189 admission date. Among hospitalizations for possible pneumonia, we also assessed the PPV when
18
19 190 radiological evidence of pneumonia was required [i.e. pneumonia, opacity, or infiltrate mentioned in a
20
21 191 chest X-ray or computed tomography scan report] (*Supplementary appendix*). All analyses were
22
23 192 performed in Stata-IC, version 15.1 (College Station TX).
24
25
26 193

28 194 **RESULTS**

30 195 *Cohort characteristics*

32 196 Among a retrospective cohort of 129,465 adults ≥ 50 years of age enrolled in TennCare, 8,322
33
34 197 hospitalizations for serious infection were identified during the study period (2008-2013). Pneumonia,
35
36 198 cellulitis and bacteremia/sepsis were the most common infections (54.3%, 20.5% and 18.4%,
37
38 199 respectively), followed by pyelonephritis (3.8%) and septic arthritis/osteomyelitis (2.5%). Fewer than 1%
39
40 200 of hospitalizations were due to meningitis/encephalitis (n=30) and endocarditis (n=18). Cohort members
41
42 201 were primarily female (57.8%) with a median age of 60 years and with residence outside of a nursing
43
44 202 home (85.9%).
45
46

47 203 *Collection, review and adjudication of selected medical records*

49 204 Of the 27 hospitals that were initially selected for the sample, 21 (78%) were able to participate.
50
51 205 We selected 7 additional hospitals to replace the 6 non-participants to achieve the desired sample size,
52
53 206 including an additional small hospital in the East region due to a large number of unavailable records
54
55 207 from a single participating hospital.
56
57

1
2
3 208 We received 716 (89%) of 808 requested records from 28 participating hospitals [Table 2].
4
5 209 Record availability from participating hospitals was lower in medium size hospitals (81.8%) compared to
6
7 210 small (93.5%) and large hospitals (91.7%), but did not differ by geographic region. Record availability by
8
9 211 infection type was greater than 86% for all infection types, with the exception of hospitalizations for the
10
11 212 rare endocarditis cases (57.1%; only 4 of 7 cases).

13
14 213 There was evidence of transfer from a prior healthcare facility for 21.8% of the hospitalizations
15
16 214 for serious infection [highest percentage of transfers for bacteremia/sepsis (38.5%) and pneumonia
17
18 215 (25.1%)]. The most common healthcare facility source was a nursing home/skilled nursing facility
19
20 216 (84.6%), but also included group home sources (7.7%), other sources (4.5%) [assisted living facility,
21
22 217 mental health center] and another acute care hospital (3.2%). There was evidence of an emergency
23
24 218 department visit within 7 days prior to admission date for the serious infection hospitalization in 4.8% of
25
26 219 the records.

28 220 *Performance of discharge diagnosis codes*

30
31 221 A total of 646 [PPV: 90.2% (95% CI: 88.0-92.4)] of the hospitalizations for serious infection
32
33 222 identified using ICD-9-CM primary discharge diagnosis codes were confirmed by applying the *a priori*
34
35 223 definitions to the abstracted data. The PPV was highest for pneumonia and cellulitis [96.8% (95% CI:
36
37 224 94.5-98.4) and 91.1% (95% CI: 86.0-96.1), respectively], and was $\geq 75\%$ for bacteremia/sepsis,
38
39 225 pyelonephritis, septic arthritis/osteomyelitis, and endocarditis. The PPV was lowest for
40
41 226 meningitis/encephalitis [50.0% (95% CI: 19.0-81.0)], although the precision was limited due to a low
42
43 227 number of available records for review (Table 2).

45 228 When performance was evaluated across stratification sampling parameters, no apparent
46
47 229 differences were observed in the PPV for records from hospitals in different geographical regions of
48
49 230 Tennessee. Although the PPV was high for all three discharge volume groups, the PPV was significantly
50
51 231 lower in large hospitals [84.6% (95% CI: 80.1-89.0)] compared to smaller hospitals [93.9% (95% CI:
52
53 232 90.8-97.0); PPV difference: -9.3% (95% CI: -14.7, -3.9)] and medium hospitals [92.7% (95% CI: 89.4-
54
55 233 96.0); PPV difference: -8.1% (95% CI: -13.7, -2.6)] (Table 2). This was likely driven by the different

234 distributions in the types of infections selected for review in the hospital groups. Large hospitals had a
235 higher proportion of non-pneumonia infections (70.4%) compared to medium and small hospitals (49.4%
236 and 36.1%, respectively). Importantly, the PPV for pneumonia was similar in each discharge volume
237 group (range: 96.0 to 96.6%), whereas the PPV was smaller for non-pneumonia infections in large
238 hospitals (79.8%) compared to medium (88.7%) and small (89.2%) hospitals.

239 In the 82 records independently abstracted by two reviewers to assess reliability, there was 92.7%
240 (95% CI: 86.9-98.4) agreement for identifying true hospitalizations for serious infection. The inter-rater
241 agreement was also high when assessing reliability, independent of the outcome prevalence, with an AC₁
242 of 0.91 (95% CI: 0.84-0.99).

243 *Sensitivity analyses*

244 The PPV was virtually unchanged when excluding the 21.8% of hospitalizations that occurred as
245 transfers from another healthcare facility [90.1% (95% CI: 87.7-92.6)]. Microbiological evidence of the
246 specific infection type was found in 47.6% of records, leading to reduced PPVs when requiring
247 microbiological evidence [45.4% (95% CI: 41.7-49.0)]. Microbiological evidence of infection was
248 highest in hospitalizations for suspected pyelonephritis (94.4%), but was $\leq 60\%$ for every other infection
249 type [pneumonia (42.7%); cellulitis/soft tissue infections (58.5%); bacteremia/sepsis (26.1%)]. When
250 requiring radiological confirmation of pneumonia, the PPV for coded diagnoses was 78.8% (95% CI:
251 74.5-83.2). Approximately 95.6% of possible hospitalizations for pneumonia had at least one documented
252 chest x-ray or CT-scan. Among those patients with a chest x-ray or CT-scan report available (n=325),
253 83.4% had a finding compatible with pneumonia. The main findings among the 54 patients with possible
254 pneumonia and a radiological report available, but without radiological confirmation of pneumonia
255 included atelectasis (n=6), interstitial pneumonitis (n=3), chronic heart failure with pulmonary edema
256 (n=1), and no radiological findings of any kind (n=44).

257

258

259

260 DISCUSSION

261 Discharge diagnoses for identifying hospitalizations due to serious infections among older adults
262 had an overall positive predictive value of 90.2%, with highest values for identification of common
263 serious infections. PPVs were consistently high across different hospital types and regions of Tennessee.
264 Microbiological confirmation was available for fewer than 50% of those admitted with possible
265 hospitalizations for serious infections, and as expected, such a requirement resulted in a lower PPV.
266 Importantly, the PPV for pneumonia hospitalizations remained relatively high even when requiring
267 radiological confirmation. In addition, including hospitalizations for serious infection that were the result
268 of a transfer from another healthcare facility (e.g. acute care hospital, skilled nursing facility) did not
269 change the PPV of hospitalizations for serious infection.

270 The PPV for hospitalizations for pneumonia in previous smaller validation studies has ranged
271 from 72 to 86% in different healthcare systems, but those studies were not focused on older adults (25,
272 41-43). In our study of hospitalizations among older adults, we found that coded discharge diagnoses
273 have a higher PPV for pneumonia compared to previous studies. The PPV for bacteremia/sepsis was also
274 on the higher range of previously reported PPVs for diagnosis codes to identify bacteremia/sepsis from
275 administrative data in other populations (reported range from 45% to 97.7%), and for septic
276 arthritis/osteomyelitis compared to a previous study conducted among patients with diabetes (63.9%
277 versus 75.9% in our study) (44-46). Overall, the observed PPV for all infections in our study was
278 comparable to two previous comprehensive validation studies of bacterial infections, one among patients
279 with rheumatoid arthritis in a single hospital system and another among patients in one of the Veteran's
280 Affairs integrated service networks (28, 34). Compared to the these two previous studies of ICD-9 codes,
281 we abstracted and adjudicated a larger number of records while using a more systematic sampling
282 strategy to retrieve and review records for hospitalizations from multiple regions and hospital types as
283 opposed to a single hospital or healthcare system. However, the PPVs for individual infections were less
284 precise and less similar to these previous studies, especially for rare infections, as would be expected due
285 to the low numbers of rare infections across previous studies (28, 34). The results of our study are also

286 similar to previous validation studies that used corresponding ICD-10 diagnosis codes to identify
287 hospitalizations for serious infection (47, 48).

288 One limitation to consider in our study was that it was not designed to estimate the sensitivity and
289 specificity of the coding algorithms. This would have required the identification, review and adjudication
290 of a sample of hospitalizations that did not fulfill our algorithm (i.e. presence of the ICD-9 primary
291 discharge diagnosis codes indicative of infection). However, when the prevalence of an outcome is low,
292 the PPV approximates the specificity (49). Importantly, any non-differential outcome misclassification
293 between exposure groups resulting from the use of imperfect but highly-specific measurements would
294 attenuate the impact of the misclassification on the relative risk estimates (50). In addition, we found that
295 the PPV of coded discharge diagnoses for serious infections remained high across hospitals of different
296 sizes and across different geographical areas of Tennessee, which may have different prevalences of
297 hospitalizations for serious infection (51). Although our study applied a systematic sampling strategy to
298 assure the representation of different settings in our population, our population was restricted to older
299 adults enrolled in a State Medicaid program. Therefore, caution is warranted when extrapolating the study
300 findings to other populations.

301 Another limitation is the use of available clinical information to operationalize definitions for
302 adjudication of true hospitalizations for infections. It is possible that some procedures, laboratory findings
303 and diagnoses that informed the final diagnosis of infection were not fully recorded in the medical
304 records, and thus, were not available for our review and may have contributed to the observed PPV for
305 some infections. Although we used previous validation studies and clinical information to build pre-
306 specified definitions for the adjudication of true infections, our reference criteria may be imperfect,
307 considering the retrospective nature of our determinations and potential variability in clinical practice.
308 Nevertheless, we also assessed how the availability of selected findings (i.e. microbiological and
309 radiological information) in the medical record impacted the overall and infection-specific PPV. We
310 demonstrated that relying on highly specific clinical diagnostics, such as microbiological and radiological
311 information, to confirm true infections would result in lower PPVs for identification of infections in

1
2
3 312 administrative data. Requiring microbiological confirmation to confirm true infections is challenging
4
5 313 because of the known low sensitivity of culture-based diagnostic methods (most commonly used in
6
7 314 clinical practice), which may lead to misclassification (52, 53). In addition, requiring radiological
8
9 315 evidence compatible with pneumonia within 2 days of hospital admission did lower the observed PPV for
10
11 316 pneumonia hospitalizations. Nevertheless, the observed PPV remained close to 80%, which should reduce
12
13 317 concerns about using diagnosis codes to identify hospitalizations due to pneumonia. Finally, the coding
14
15 318 algorithms were based on the ICD-9-coding system only. Although these findings will be helpful for
16
17 319 retrospective studies that encompass periods of ICD-9 use, additional studies evaluating the performance
18
19 320 of ICD-10-based codes would be useful to complement our findings.
20
21

22 321 Our study demonstrated that discharge diagnosis codes can be used to accurately identify
23
24 322 hospitalizations for serious infections among older adults. The highest PPVs were observed for the most
25
26 323 common infections, and the PPV for pneumonia remained high when requiring radiological confirmation.
27
28 324 The PPV was poor when microbiological confirmation of infection was required to identify a true
29
30 325 hospitalization for serious infection. This information supports the use of discharge diagnosis codes for
31
32 326 infections as outcomes in ongoing and future studies among older adults.
33
34

35 327

36 328 **Acknowledgement**

37
38
39 329 We are indebted to the Tennessee Bureau of TennCare of the Department of Finance and Administration,
40
41 330 which provided data for the study. We are also indebted to the Tennessee Department of Health for
42
43 331 providing data for the study. Statements in the report should not be construed as endorsement by the U.S.
44
45 332 Department of Health and Human Services, the Department of Veterans Affairs, or the Tennessee
46
47 333 Department of Health.
48

49 334

50 335 **Funding**

51
52
53 336 This study was funded by the NIH (R03-AG-042981 and R01-AG-043471-01A1) and the TL1 award
54
55 337 TL1TR000447.
56

1
2
3 338
4
5 339 **Contributors**
6
7 340 ADW planned the medical record collection and statistical analysis, analyzed and interpreted the data, and
8
9 341 drafted and revised the paper. MRG, WS, CMS, and RAG planned the statistical analysis, interpreted the
10
11 342 data and revised the paper. EFM prepared the data, and revised the paper. CGG initiated the project,
12
13 343 acquired the data from TennCare, planned the medical record collection and statistical analysis,
14
15 344 interpreted the data, and revised the paper.
16
17
18 345

19
20 346 **Declaration of interests**
21

22 347 CGG has received consulting fees from Pfizer and Merck, and received research support from Sanofi-
23
24 348 Pasteur, Campbell Alliance, the Centers for Disease Control and Prevention, National Institutes of Health,
25
26 349 and the Agency for Health Care Research and Quality. WS has received personal fees from Pfizer,
27
28 350 Merck, Novavax, Dynavax, Sanofi-Pasteur, GSK, Seqirus, and received research support from the
29
30 351 Centers for Disease Control and Prevention. ADW, EFM, CMS, MRG and RAG have no conflicts of
31
32 352 interest to disclose.
33
34
35 353

36
37 354 **Data sharing**
38

39 355 No additional unpublished data are available from the study. The study protocol and statistical code are
40
41 356 available from the corresponding author, Andrew Wiese (andrew.d.wiese@vanderbilt.edu).
42
43 357
44
45 358
46
47 359
48
49 360
50
51 361
52
53 362
54
55 363
56
57

REFERENCES

1. Khabbaz RF, Moseley RR, Steiner RJ, Levitt AM, Bell BP. Challenges of infectious diseases in the USA. *Lancet* 2014; **384**(9937): 53-63.
2. Grohskopf LA, Sokolow LZ, Broder KR, et al. Prevention and Control of Seasonal Influenza with Vaccines. *MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control* 2016; **65**(5): 1-54.
3. Kim DK, Bridges CB, Harriman KH. Advisory committee on immunization practices recommended immunization schedule for adults aged 19 years or older--United States, 2015. *MMWR Morbidity and mortality weekly report* 2015; **64**(4): 91-2.
4. Jeon CY, Muennig P, Furuya EY, Cohen B, Nash D, Larson EL. Burden of present-on-admission infections and health care-associated infections, by race and ethnicity. *American journal of infection control* 2014; **42**(12): 1296-302.
5. Pilishvili T, Bennett NM. Pneumococcal Disease Prevention Among Adults: Strategies for the Use of Pneumococcal Vaccines. *American journal of preventive medicine* 2015; **49**(6 Suppl 4): S383-90.
6. Sjoding MW, Prescott HC, Wunsch H, Iwashyna TJ, Cooke CR. Longitudinal Changes in ICU Admissions Among Elderly Patients in the United States. *Critical care medicine* 2016; **44**(7): 1353-60.
7. Crotty MP, Meyers S, Hampton N, et al. Epidemiology, Co-Infections, and Outcomes of Viral Pneumonia in Adults: An Observational Cohort Study. *Medicine* 2015; **94**(50): e2332.
8. Simonetti AF, Garcia-Vidal C, Viasus D, et al. Declining mortality among hospitalized patients with community-acquired pneumonia. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2016; **22**(6): 567.e1-7.
9. Rivero-Calle I, Pardo-Seco J, Aldaz P, et al. Incidence and risk factor prevalence of community-acquired pneumonia in adults in primary care in Spain (NEUMO-ES-RISK project). *BMC infectious diseases* 2016; **16**(1): 645.
10. Park JY, Park S, Lee SH, et al. Microorganisms Causing Community-Acquired Acute Bronchitis: The Role of Bacterial Infection. *PloS one* 2016; **11**(10): e0165553.

- 1
2
3 390 11. Jackson ML, Walker R, Lee S, Larson E, Dublin S. Predicting 2-Year Risk of Developing
4
5 391 Pneumonia in Older Adults without Dementia. *Journal of the American Geriatrics Society* 2016; **64**(7):
6
7 392 1439-47.
8
9 393 12. Clement RC, Haddix KP, Creighton RA, Spang JT, Tennant JN, Kamath GV. Risk Factors for
10
11 394 Infection After Knee Arthroscopy: Analysis of 595,083 Cases From 3 United States Databases.
12
13 395 *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy*
14
15 396 *Association of North America and the International Arthroscopy Association* 2016; **32**(12): 2556-61.
16
17 397 13. Weinert BA, Edmonson MB. Hospitalizations at Nonfederal Facilities for Lower Respiratory
18
19 398 Tract Infection in American Indian and Alaska Native Children Younger than 5 Years of Age, 1997-2012.
20
21 399 *The Journal of pediatrics* 2016; **175**: 33-9.e4.
22
23
24 400 14. Akiyama T, Chikuda H, Yasunaga H, Horiguchi H, Fushimi K, Saita K. Incidence and risk
25
26 401 factors for mortality of vertebral osteomyelitis: a retrospective analysis using the Japanese diagnosis
27
28 402 procedure combination database. *BMJ open* 2013; **3**(3).
29
30 403 15. Graversen ME, Dalgaard LS, Jensen-Fangel S, Jespersen B, Ostergaard L, Sogaard OS. Risk and
31
32 404 outcome of pyelonephritis among renal transplant recipients. *BMC infectious diseases* 2016; **16**: 264.
33
34 405 16. Abou Zahr Z, Spiegelman A, Cantu M, Ng B. Perioperative use of anti-rheumatic agents does not
35
36 406 increase early postoperative infection risks: a Veteran Affairs' administrative database study.
37
38 407 *Rheumatology international* 2015; **35**(2): 265-72.
39
40 408 17. Carnahan RM, Moores KG, Perencevich EN. A systematic review of validated methods for
41
42 409 identifying infection related to blood products, tissue grafts, or organ transplants using administrative
43
44 410 data. *Pharmacoepidemiology and drug safety* 2012; **21 Suppl 1**: 213-21.
45
46 411 18. Condell O, Gubbels S, Nielsen J, et al. Automated surveillance system for hospital-acquired
47
48 412 urinary tract infections in Denmark. *The Journal of hospital infection* 2016; **93**(3): 290-6.
49
50 413 19. Gedeberg R, Furebring M, Michaelsson K. Diagnosis-dependent misclassification of infections
51
52 414 using administrative data variably affected incidence and mortality estimates in ICU patients. *Journal of*
53
54 415 *clinical epidemiology* 2007; **60**(2): 155-62.
55
56
57
58
59
60

- 1
2
3 416 20. Goto M, Ohl ME, Schweizer ML, Perencevich EN. Accuracy of administrative code data for the
4
5 417 surveillance of healthcare-associated infections: a systematic review and meta-analysis. *Clinical*
6
7 418 *infectious diseases : an official publication of the Infectious Diseases Society of America* 2014; **58**(5):
8
9 419 688-96.
- 11 420 21. Shaklee J, Zerr DM, Elward A, et al. Improving surveillance for pediatric *Clostridium difficile*
12
13 421 infection: derivation and validation of an accurate case-finding tool. *The Pediatric infectious disease*
14
15 422 *journal* 2011; **30**(3): e38-40.
- 18 423 22. Wright SB, Huskins WC, Dokholyan RS, Goldmann DA, Platt R. Administrative databases
19
20 424 provide inaccurate data for surveillance of long-term central venous catheter-associated infections.
21
22 425 *Infection control and hospital epidemiology : the official journal of the Society of Hospital*
23
24 426 *Epidemiologists of America* 2003; **24**(12): 946-9.
- 26 427 23. Greenberg JA, Hohmann SF, Hall JB, Kress JP, David MZ. Validation of a Method to Identify
27
28 428 Immunocompromised Patients with Severe Sepsis in Administrative Databases. *Annals of the American*
29
30 429 *Thoracic Society* 2016; **13**(2): 253-8.
- 32 430 24. Olsen MA, Young-Xu Y, Stwalley D, et al. The burden of *clostridium difficile* infection:
33
34 431 estimates of the incidence of CDI from U.S. Administrative databases. *BMC infectious diseases* 2016; **16**:
35
36 432 177.
- 39 433 25. Grijalva CG, Chung CP, Stein CM, et al. Computerized definitions showed high positive
40
41 434 predictive values for identifying hospitalizations for congestive heart failure and selected infections in
42
43 435 Medicaid enrollees with rheumatoid arthritis. *Pharmacoepidemiology and drug safety* 2008; **17**(9): 890-5.
- 45 436 26. TennCare Fiscal Year 2015-2016 Annual Report. In: Administration DoHCF, ed.2016.
- 47 437 267. Wiese AD, Griffin MR, Stein CM, Mitchel EF, Jr., Grijalva CG. Opioid Analgesics and the Risk
48
49 438 of Serious Infections Among Patients With Rheumatoid Arthritis: A Self-Controlled Case Series Study.
50
51 439 *Arthritis & rheumatology (Hoboken, NJ)* 2016; **68**(2): 323-31.

- 1
2
3 440 28. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital
4
5 441 discharge databases coded serious bacterial infections accurately. *Journal of clinical epidemiology* 2007;
6
7 442 **60**(4): 397-409.
8
9 443 29. Grijalva CG, Kaltenbach L, Arbogast PG, Mitchel EF, Jr., Griffin MR. Initiation of rheumatoid
10
11 444 arthritis treatments and the risk of serious infections. *Rheumatology (Oxford, England)* 2010; **49**(1): 82-
12
13 445 90.
14
15 446 30. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values
16
17 447 with disease prevalence. *Statistics in medicine* 1997; **16**(9): 981-91.
18
19 448 31. Kim HM, Smith EG, Stano CM, et al. Validation of key behaviourally based mental health
20
21 449 diagnoses in administrative data: suicide attempt, alcohol abuse, illicit drug abuse and tobacco use. *BMC*
22
23 450 *health services research* 2012; **12**: 18.
24
25 451 32. El-Ghitany EM, Farghaly AG, Farag S, Abd El-Wahab EW. Validation of EGCRISC for Chronic
26
27 452 Hepatitis C Infection Screening and Risk Assessment in the Egyptian Population. *PloS one* 2016; **11**(12):
28
29 453 e0168649.
30
31 454 33. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data
32
33 455 capture (REDCap)--a metadata-driven methodology and workflow process for providing translational
34
35 456 research informatics support. *Journal of biomedical informatics* 2009; **42**(2): 377-81.
36
37 457 34. Patkar NM, Curtis JR, Teng GG, et al. Administrative codes combined with medical records
38
39 458 based criteria accurately identified bacterial infections among rheumatoid arthritis patients. *Journal of*
40
41 459 *clinical epidemiology* 2009; **62**(3): 321-7, 7.e1-7.
42
43 460 35. Aronson PL, Williams DJ, Thurm C, et al. Accuracy of diagnosis codes to identify febrile young
44
45 461 infants using administrative data. *Journal of hospital medicine : an official publication of the Society of*
46
47 462 *Hospital Medicine* 2015; **10**(12): 787-93.
48
49 463 36. Govatsmark RE, Sneeggen S, Karlsaune H, Slordahl SA, Bonna KH. Interrater reliability of a
50
51 464 national acute myocardial infarction register. *Clinical epidemiology* 2016; **8**: 305-12.
52
53
54
55
56
57
58
59
60

- 1
2
3 465 37. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and
4
5 466 Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality
6
7 467 disorder samples. *BMC medical research methodology* 2013; **13**: 61.
8
9 468 38. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement.
10
11 469 *The British journal of mathematical and statistical psychology* 2008; **61**(Pt 1): 29-48.
12
13 470 39. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: resolution of the two
14
15 471 paradoxes. *Journal of clinical epidemiology* 1996; **49**(4): 431-4.
16
17 472 40. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes.
18
19 473 *Journal of clinical epidemiology* 1990; **43**(6): 543-9.
20
21 474 41. Drahos J, Vanwormer JJ, Greenlee RT, Landgren O, Koshiol J. Accuracy of ICD-9-CM codes in
22
23 475 identifying infections of pneumonia and herpes simplex virus in administrative data. *Annals of*
24
25 476 *epidemiology* 2013; **23**(5): 291-3.
26
27 477 42. Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients
28
29 478 with pneumonia. *American journal of medical quality : the official journal of the American College of*
30
31 479 *Medical Quality* 2005; **20**(6): 319-28.
32
33 480 43. Kern DM, Davis J, Williams SA, et al. Validation of an administrative claims-based diagnostic
34
35 481 code for pneumonia in a US-based commercially insured COPD population. *International journal of*
36
37 482 *chronic obstructive pulmonary disease* 2015; **10**: 1417-25.
38
39 483 44. Iwashyna TJ, Odden A, Rohde J, et al. Identifying patients with severe sepsis using
40
41 484 administrative claims: patient-level validation of the angus implementation of the international consensus
42
43 485 conference definition of severe sepsis. *Medical care* 2014; **52**(6): e39-43.
44
45 486 45. Newton KM, Wagner EH, Ramsey SD, et al. The use of automated data to identify complications
46
47 487 and comorbidities of diabetes: a validation study. *Journal of clinical epidemiology* 1999; **52**(3): 199-207.
48
49 488 46. Carnahan RM, Herman RA, Moores KG. A systematic review of validated methods for
50
51 489 identifying transfusion-related sepsis using administrative and claims data. *Pharmacoepidemiology and*
52
53 490 *drug safety* 2012; **21 Suppl 1**: 222-9.

- 1
2
3 491 47. Sahli L, Lapeyre-Mestre M, Derumeaux H, Moulis G. Positive predictive values of selected
4
5 492 hospital discharge diagnoses to identify infections responsible for hospitalization in the French national
6
7 493 hospital database. *Pharmacoepidemiology and drug safety* 2016; **25**(7): 785-9.
8
9 494 48. Holland-Bill L, Xu H, Sorensen HT, et al. Positive predictive value of primary inpatient discharge
10
11 495 diagnoses of infection among cancer patients in the Danish National Registry of Patients. *Annals of*
12
13 496 *epidemiology* 2014; **24**(8): 593-7, 7.e1-18.
14
15 497 49. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic
16
17 498 research on therapeutics. *Journal of clinical epidemiology* 2005; **58**(4): 323-37.
18
19 499 50. Rodgers A, MacMahon S. Systematic underestimation of treatment effects as a result of
20
21 500 diagnostic test inaccuracy: implications for the interpretation and design of thromboprophylaxis trials.
22
23 501 *Thrombosis and haemostasis* 1995; **73**(2): 167-71.
24
25 502 51. van Walraven C, English S, Austin PC. Administrative database code accuracy did not vary
26
27 503 notably with changes in disease prevalence. *Journal of clinical epidemiology* 2016; **79**: 86-9.
28
29 504 52. Mancini N, Carletti S, Ghidoli N, Cichero P, Burioni R, Clementi M. The era of molecular and
30
31 505 other non-culture-based methods in diagnosis of sepsis. *Clinical microbiology reviews* 2010; **23**(1): 235-
32
33 506 51.
34
35 507 53. Niederman MS. The argument against using quantitative cultures in clinical trials and for the
36
37 508 management of ventilator-associated pneumonia. *Clinical infectious diseases : an official publication of*
38
39 509 *the Infectious Diseases Society of America* 2010; **51** Suppl 1: S93-9.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

516 **Table 1. Discharge diagnosis code definitions (ICD-9-CM) for hospitalizations for serious infection**

Serious Infection	Primary (first listed) discharge diagnosis code
Pneumonia-primary definition	003.22, 480.* [†] , 481, 482.*, 483.*, 484.*, 485.*, 486.*, 487.0
Pneumonia-secondary definition (pneumonia diagnosis (above) in any other diagnosis field)	510.*, 038.*, 790.7, 995.91, 995.92
Meningitis/ Encephalitis	003.21, 036.0, 047.*, 049.*, 053.0, 054.72, 072.1, 091.81, 094.2, 098.82, 100.81, 320.*, 036.1, 054.3, 056.01, 058.21, 058.29, 062.*, 063.*, 064.*, 066.41, 072.2, 094.81, 130.0, 323.*
Bacteremia/ Sepsis [†]	038.*, 790.7, 995.91, 995.92
Cellulitis/ Soft-tissue infections	035, 040.0, 569.61, 681.*, 682.*, 728.86, 785.4
Endocarditis	036.42, 074.22, 093.2*, 098.84, 421.*, 422.92
Pyelonephritis	590.*
Septic Arthritis/ Osteomyelitis	003.23, 056.71, 098.5*, 711.0, 711.00-711.07, 711.09, 711.9*, 003.24, 376.03, 526.4, 730.0*, 730.1*, 730.2*

517 † Without a diagnosis of pneumonia in any other diagnosis field

518 ‡ A * indicates all numeric values [0-9]

519

Table 2. Positive predictive value (PPV) of coded discharge diagnosis definitions for hospitalizations for serious infections among older adults enrolled in Tennessee Medicaid, 2008-2013

Type	Expected Number of Records	Records Received	PPV (95 % CI)
Overall	675	716	90.2 (88.0, 92.4)
Region Specific			
West	225	195	91.3 (87.3, 95.2)
Central	225	225	88.9 (84.8, 93.0)
East	225	296	90.5 (87.2, 93.9)
Bed volume size specific			
Low	225	230	93.9 (90.8, 97.0)
Medium	225	233	92.7 (89.4, 96.0)
High	225	253	84.6 (80.1, 89.0)
Serious Infection			
Pneumonia	305	340	96.8 (94.5, 98.4)
Cellulitis/Soft-tissue infections	125	123	91.1 (86.0, 96.1)
Pyelonephritis	80	89	87.6 (80.8, 94.5)
Bacteremia/Sepsis	100	92	82.6 (74.9, 90.4)
Septic Arthritis/Osteomyelitis	50	58	75.9 (64.8, 86.9)
Meningitis/Encephalitis	10	10	50.0 (19.0, 81.0)
Endocarditis	5	4	75.0 (32.6, 100.0)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure legends

Figure 1. Sampling strategy for identifying potential hospitalizations for serious infection

For peer review only

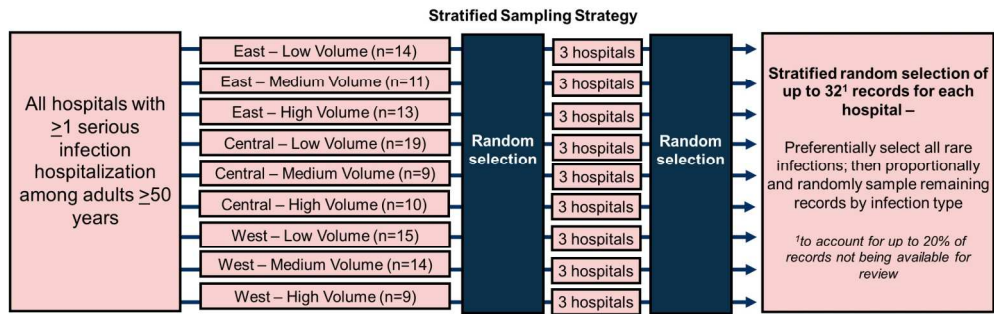


Figure 1. Sampling strategy for identifying potential hospitalizations for serious infection

271x90mm (300 x 300 DPI)

peer review only

Supplementary Appendix

Infection-Specific Definitions of Hospitalization for Serious Infection

We used a pre-specified adjudication process to determine whether each abstracted medical record corresponded to a true infection or not. Previous validation studies and expert clinical knowledge were used to define specific a priori definitions for each infection type.¹⁻³ Information abstracted from the medical record was compared to these a priori definitions for each infection type to make the final determination of whether a hospitalization represented a true infection or not.

Outline

I. Sepsis/Septicemia/Bacteremia/Septic Shock/Generalized Infection	Page 2
II. Pneumonia	Page 3
III. Cellulitis/Soft-tissue infections	Page 5
IV. Endocarditis.....	Page 6
V. Meningitis/Encephalitis	Page 7
VI. Pyelonephritis.....	Page 9
VII. Septic Arthritis/Osteomyelitis	Page 10
VIII. References.....	Page 11

I. Sepsis/Septicemia/Bacteremia/Septic Shock/Generalized Infection

Either of the following [1 or 2]:

1. Positive culture of a non-contaminant pathogen

i. Positive blood culture [any of the following (1-2)]

1. Any gram-negative organism, except:
 - a. No predominant organism
2. A gram positive organism, except:
 - a. Coagulase-negative *Staphylococcus*
 - b. *Bacillus spp.* (other than *Bacillus anthracis*)
 - c. *Corynebacterium spp.*
 - d. *Propionibacterium spp.*
 - e. *Micrococcus*
 - f. Diptheroids
 - g. Viridians Group Streptococci
 - h. Enterococci
 - i. *Clostridium perfringens*
 - j. *Aerococcus*
 - k. *Alcaligenes faecalis*
 - l. *Citrobacter*
 - m. *Neisseria subflava*
 - n. *Stomatococcus*
 - o. *Streptococcus bovis*
 - p. *Veillonella candidemia*
 - q. *Mycobacterium tuberculosis*
 - r. *S. salivarius*
 - s. "Gram Positive"
 - t. "No predominant organism"
 - u. *Streptococcus alpha*

2. At least two of the following, documented at admission +/- 2 days [i-iii]

- i. Hypotension
 1. Systolic BP \leq 90 mmHg
 2. Reduction of systolic BP of 40mmHg from earliest measurement collected during the admission of interest
- ii. Two of the following [1-4]:
 1. Temperature \geq 38⁰C **or** \leq 36⁰C
 2. Heart rate \geq 90 beats/minute
 3. Respiratory rate \geq 20 breaths/min or PaCO₂ < 32 mmHg
 4. WBC \geq 10,000 cells/mm³ **or** \leq 4,500 cells/mm³ **or** WBC with > 10 % immature (band) forms
- iii. Initiation of antibiotic treatment specifically for sepsis/septicemia/bacteremia/septic shock/generalized infection

II. Pneumonia

1. Pneumonia identified through examination (**all three of the following [a-c]**):
 - a. One of the following admission findings indicative of respiratory findings:
 1. New and/or increased cough
 2. Shortness of breath
 3. Pleuritic chest pain
 4. New purulent production
 5. Altered mental status (“agitation” and “lethargy” included)
 6. Crackles
 - a. Physical evidence of consolidation such as egophony, whispered pectoriloquy, etc.
 - b. One of the following examination findings indicative of systemic infection [1-4]:
 1. Temperature ($T \geq 100.4^{\circ}\text{F}$ (38°C) or $\leq 96^{\circ}\text{F}$) in first 48 hours of admission
 2. Systolic BP $\leq 90\text{mmHg}$
 3. Shock
 - a. Volume nonresponsive hypotension
 4. Blood peripheral WBC ($\geq 10.0 \times 10^9/\text{L}$ or $\leq 4.5 \times 10^9/\text{L}$)
 - c. Treatment with antibiotics/antivirals indicated for suspected infection

OR

At least two of the following [1-3]:

1. Two of the following from #1 (**[a and b], [a and c], or [b-c]**)
2. Any of the following findings listed on chest imaging from radiologic report **documented at admission +/- 2 days**
 - a. Pneumonia
 - b. Lung abscess
 - c. Opacity consistent with pneumonia/lung abscess
 - d. Infiltrate consistent with pneumonia/lung abscess
 - e. Consolidation consistent with pneumonia/lung abscess
 - f. Increased density consistent with pneumonia/lung abscess
 - g. Pleural effusion consistent with pneumonia/lung abscess
 - h. Interstitial edema consistent with pneumonia/lung abscess
3. Sterile Site Laboratory Findings
 - i. Any one of the following [i through v]
 - i. Sputum lab findings [any **one** of the following (1, 2)]:
 1. Sputum culture/PCR/serology/gram stain positive for an agent that is not considered a contaminant [see exclusion list below]:
 - a. *Aspergillus* species, *Enterococcus* species, viridians group streptococci, and yeast
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - ii. Blood lab findings [either of the following (1-3)]
 1. Blood culture/PCR/serology positive for an agent that is not considered a contaminant [see exclusion list below]:
 - a. Exclusions
 - i. Coagulase-negative *Staphylococcus*

- ii. *Bacillus spp.* (other than *Bacillus anthracis*)
 - iii. *Corynebacterium spp.*
 - iv. *Propionibacterium spp.*
 - v. *Micrococcus*
 - vi. Diphtheroids
 - vii. Viridians Group Streptococci
 - viii. Enterococci
 - ix. *Clostridium perfringens*
 - x. *Aerococcus*
 - xi. *Alcaligenes faecalis*
 - xii. *Citrobacter*
 - xiii. *Neisseria subflava*
 - xiv. *Stomatococcus*
 - xv. *Streptococcus bovis*
 - xvi. *Veillonella candidemia*
 - xvii. *Mycobacterium tuberculosis*
 - xviii. *S. salivarius*
2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
- iii. Pleural fluid lab findings [either of the following (1, 2)]
 1. Culture/PCR/serology positive for a bacterial pathogen
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - iv. Bronchoscopic specimen or deep endotracheal tube aspiration lab findings [either of the following (1, 2)]
 1. Culture/PCR/serology positive for a bacterial pathogen
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - v. Urine antigen detection testing [either of the following (1, 2)]
 1. *Legionella pneumophila*
 2. *Streptococcus pneumoniae*

III. Cellulitis/Soft-Tissue Infection

Both of the following:

1. Any mention of the following with recent onset (≤ 14 days) [*any of the following*]
 - a. Skin erythema
 - b. Surgical site infection
 - c. Superficial central line infection
 - d. Ostomy site infection
 - e. Skin infection with associated lymphangitis
2. Antibiotic treatment initiated for suspected infection

For peer review only

IV. Endocarditis

Any one of the following [1-3]:

1. Major Criteria [both of the following]:
 - a. Suggestive microbiology [at least one of the following]:
 - i. Positive blood culture of an **endocarditis organism** [any of the following]:
 1. *Streptococcus bovis*
 2. *Viridians streptococci*
 3. *Staphylococcus aureus*
 4. *Enterococcus spp.*
 5. HACEK organisms
 6. Coagulase negative staphylococci
 - b. Evidence of endocardial involvement [at least one of the following]:
 - i. New regurgitant murmur (a change in a preexisting murmur does not get scored)
 - ii. Echocardiogram suspicious for any of the following:
 1. Intracardiac mass with no alternative explanation
 2. Endocardial abscess
 3. New partial prosthesis dehiscence
 4. Vegetation on valve
2. Minor Criteria [at least 4 of the following]:
 - a. Predisposing valvular disease or IV drug use
 - b. Temperature $\geq 100.4^{\circ}\text{F}$ or 38°C
 - c. Vascular phenomena
 - i. Janeway lesions, conjunctival hemorrhages, arterial emboli, septic pulmonary infarcts, mycotic aneurysm, intracranial bleed
 - d. Immunologic phenomena
 - i. Osler nodes, Roth Spots, elevated Rheumatoid factor, hematuria in non-catheter urine, or other evidence of glomerulonephritis
 - e. Positive blood cultures
 - i. Excluding a single positive culture for coagulase negative staphylococci or a single positive culture for an organism that does not fall into the “reasonable endocarditis organism” (i.e. coagulase-positive and coagulase-negative *S. aureus*, Enterococcus, viridians group Streptococci, *S. bovis*, HACEK organisms)
 - f. Positive serology for Brucella, Bartonella, Legionella, or Chlamydia
 - g. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
3. At least one Major Criteria **AND** 3 minor criteria.

V. Meningitis/Encephalitis

Any one of the following [1 or 2]:

1. Both of the following [a-b]
 - a. Laboratory Findings [any one of the following (i-ix)]
 - i. CSF demonstrates any bacterium
 1. Excluding Diptheroids, Propionibacteria, Bacillus, Coagulase Negative *Staphylococcus*
 - ii. CSF demonstrates Diptheroids, Propionibacteria, Bacillus, Coagulase Negative *Staphylococcus* in the setting of past neurosurgical intervention **AND** physicians elected to treat with antibacterials
 - iii. Blood cultures positive for any of the following:
 1. *S. pneumoniae*
 2. *H. influenza*
 3. *Neisseria meningitidis*
 4. Group B Streptococcus
 - iv. Stool cultures positive for enterovirus
 - v. Throat or sputum cultures positive for *Neisseria meningitidis* in the setting of a rapid onset, overwhelming infection syndrome, including petechiae
 - vi. Serology positive for *Mycoplasma*, *Leptospira*, measles, mumps, lymphocytic choriomeningitis virus, arboviruses (e.g. St. Louis encephalitis virus), or HIV (if historically consistent with acute seroconversion).
 - vii. Brain biopsy demonstrates encephalitis
 - viii. Positive CSF culture or PCR detection for any of the following
 - ix. Acute or convalescent serology demonstrates positive antibody pattern for any of the following:
 1. Encephalitis arbovirus (La Crosse, St. Louis, Eastern Equine, Western Equine, Powassan, Japanese, West Nile)
 - b. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected meningitis/encephalitis
2. At least two of the following [a-d]
 - a. Clinical meningitis/encephalitis [at least two of the following]:
 - i. Petechial rash
 - ii. Nuchal rigidity (by history or exam)
 - iii. Altered sensorium
 - iv. Fever
 - v. Altered level of consciousness, including “agitation” or “lethargy”
 - vi. Behavioral change
 - vii. Diminished level of consciousness (not easily roused)
 - viii. History of any of the following: headaches, altered mental status, or recent exposure to patient with known bacterial meningitis
 - ix. Reduction in fever within 72 hours of starting anti-bacterial
 - b. Inflammatory CSF [at least one of the following i-ii]
 - i. Pleocytosis: ≥ 15 WBC/mm³ (after subtracting one WBC for every 1,000 RBC)
 - ii. Elevated protein (based on local lab-determined upper limits)
 - c. Suggestive Findings [at least one of the following (i-iv)]

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
- i. Septic syndrome
 - ii. Focal neurological deficits documented during examination (such as flaccid paralysis or speech alterations for West Nile Virus)
 - iii. Abnormal imaging
 1. Computed tomography or magnetic resonance imaging (MRI) demonstrating focal edema or inflammation or hemorrhage
 2. Indicated as “meningitis/encephalitis” or “compatible with meningitis/encephalitis” or “cannot rule out meningitis/encephalitis”
 - iv. Findings indicating an abnormal electroencephalography (such as focal periodic discharges)
- d. Antibiotic/antiviral/antifungal treatment initiated/recommended for presumed meningitis/encephalitis

VI. Pyelonephritis

At least two of the following [1-4]:

1. Suggestion of infection [at least one of the following]:
 - a. Temperature \geq 100.4°F (38°C)
 - b. Peripheral blood WBC \geq 10,000/mm³
 - c. Positive blood culture for any of the following:
 - i. Gram Negative Rods
 - ii. *Enterococcus spp.*
 - iii. *Staphylococcus saprophyticus*
 - d. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
2. Strong renal localization [at least one of the following]:
 - a. CT, MRI, or Ultrasound Suggestive of Renal Inflammation
3. Minor Criteria [at least two of the following]:
 - a. Flank pain
 - b. Costovertebral angle tenderness
 - c. Complaints of dysuria, frequency, or suprapubic pain
 - d. Any pyuria
 - e. Urine culture positive for a single organism
4. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected pyelonephritis

VII. Septic Arthritis/Osteomyelitis

Any one of the following (1-5):

1. Synovial fluid gram stain or tissue gram stain or special stain demonstrating any organism
2. Joint culture/PCR/serology positive for any organism
3. At least two of the following (a-d):
 - a. Positive blood culture/PCR/serology
 - b. Joint with acute (≤ 7 days) worsening of inflammatory features (**at least two of the following**):
 - i. Pain on history
 - ii. ROM
 - iii. Warmth
 - iv. Effusion
 - v. Swelling
 - vi. Limited range of motion
 - c. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
 - d. Any one of the following (i-iv)
 - i. Synovial fluid WBC $\geq 30,000/\text{mm}^3$
 - ii. Synovial fluid WBC $\geq 60,000/\text{mm}^3$ with $> 75\%$ PMNs
 - iii. Skin lesions, tenosynovitis, or urethral/cervical/rectal Gram stain or culture suggestive of *Neisseria gonorrhoeae*
 - iv. Any indication of the following in the synovial fluid: needle-like crystals, CPPD crystals, uric acid.
4. Positive bone biopsy [at least one of the following (a-c)]:
 - a. Positive culture for any organism
 - b. Positive gram stain
5. Imaging and indirect features [**at least two of the following (a-c)**]:
 - a. Consistent imaging [at least one of the following (i-iv)]:
 - i. Plain X-ray read by a radiologist as suggestive of osteomyelitis
 - ii. CT Scan read by a radiologist as suggestive of osteomyelitis
 - iii. MRI read by a radiologist as suggestive of osteomyelitis
 - iv. Bone scan or WBC scan read as suggestive of osteomyelitis
 - b. Suggestive indirect features[at least one of the following (i-viii)]:
 - i. Temperature $> 100.4^\circ\text{F}$ (38°C)
 - ii. Bony pain or tenderness or erythema over bone suspected to be infected
 - iii. Draining soft tissue sinus over bone suspected to be infected
 - iv. Positive “probe to bone” (or visible bone in deep ulcer at suspected site)
 - v. Blood culture positive for *S. aureus*
 - vi. ESR ≥ 75 mm/hour
 - vii. Intravenous drug use or indwelling catheter
 - viii. Inflammation on imaging associated with an orthopedic prosthesis
 - c. Positive culture for any organism from wound sample over the bone suspected of infection
 - d. Antibiotic/antiviral/antifungal treatment for suspected infection

VIII. References

1. Grijalva CG, Chung CP, Stein CM, et al. Computerized definitions showed high positive predictive values for identifying hospitalizations for congestive heart failure and selected infections in Medicaid enrollees with rheumatoid arthritis. *Pharmacoepidemiology and drug safety* 2008; 17(9): 890-5.
2. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital discharge databases coded serious bacterial infections accurately. *Journal of clinical epidemiology* 2007; 60(4): 397-409.
3. Patkar NM, Curtis JR, Teng GG, et al. Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients. *Journal of clinical epidemiology* 2009; 62(3): 321-7, 7.e1-7.

Section & Topic	No	Item	Reported on page #
Title or Abstract			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
<i>Participants</i>	6	Eligibility criteria	4,5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4,5
	9	Whether participants formed a consecutive, random or convenience series	4-6
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5, Table 1, Supplementary Appendix
	10b	Reference standard, in sufficient detail to allow replication	6, Supplementary Appendix
	11	Rationale for choosing the reference standard (if alternatives exist)	6, Supplementary Appendix
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6, Supplementary Appendix
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6, Supplementary Appendix
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6, Supplementary Appendix
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	7,8
	15	How indeterminate index test or reference standard results were handled	7,8
	16	How missing data on the index test and reference standard were handled	7,8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7,8
	18	Intended sample size and how it was determined	7,8
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	8
	20	Baseline demographic and clinical characteristics of participants	8
	21a	Distribution of severity of disease in those with the target condition	n/a
	21b	Distribution of alternative diagnoses in those without the target condition	n/a
	22	Time interval and any clinical interventions between index test and reference standard	n/a
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	9,10, Table 2
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	9,10 Table 2
	25	Any adverse events from performing the index test or the reference standard	n/a
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	11-13
	27	Implications for practice, including the intended use and clinical role of the index test	11-13

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

OTHER INFORMATION			
	28	Registration number and name of registry	n/a
	29	Where the full study protocol can be accessed	14
	30	Sources of funding and other support; role of funders	13

For peer review only

BMJ Open

Validation of discharge diagnosis codes to identify serious infections among middle age and older adults

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020857.R1
Article Type:	Research
Date Submitted by the Author:	14-Mar-2018
Complete List of Authors:	Wiese, Andrew; Vanderbilt University Medical Center, Health Policy Griffin, Marie R; Vanderbilt University Medical Center, Health Policy Stein, Michael; Vanderbilt University, Pharmacology Schaffner, William; Vanderbilt University Medical Center, Health Policy Greevy, Robert; Vanderbilt School of Medicine, Biostatistics Mitchel, Jr., Edward; Vanderbilt University Medical Center, Health Policy Grijalva, Carlos; Vanderbilt University, Health Policy
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Infectious diseases, Epidemiology
Keywords:	coding algorithms, Medicaid, older adults, serious infections

SCHOLARONE™
Manuscripts

Peer Review Only

1
2
3 1 **Validation of discharge diagnosis codes to identify serious infections among middle age and older**
4
5 2 **adults**
6

7 3
8
9 4 **Running title:** Validation of diagnosis codes to identify infections
10
11 5

12
13 6 **Authors:** Andrew D. Wiese, PhD, MPH¹; Marie R. Griffin^{1,2}, MD, MPH; C. Michael Stein, MB, ChB³;
14 7 William Schaffner, MD¹; Robert Greevy, PhD⁴; Edward F. Mitchel Jr., MS¹; Carlos G. Grijalva, MD,
15 8 MPH^{1,2}
16
17

18
19 9 **Affiliations:** ¹Department of Health Policy, Vanderbilt University School of Medicine, Nashville,
20 10 Tennessee, USA; ²Mid-South Geriatric Research Education and Clinical Center, VA Tennessee Valley
21 11 Health Care System, Nashville, Tennessee, USA; ³Departments of Pharmacology and ⁴Biostatistics,
22 12 Vanderbilt University School of Medicine, Nashville, Tennessee, USA
23
24

25 13 **Corresponding Author:** Andrew D. Wiese, PhD, MPH; Department of Health Policy, Vanderbilt
26 14 University Medical Center, Suite 2600, Village at Vanderbilt, 1500 21st Avenue South, Nashville, TN
27 15 37212; phone: (615) 875-7997; email: andrew.d.wiese@vanderbilt.edu
28
29

30 16
31 17 **Key words:** coding algorithms; Medicaid; older adults; serious infections
32
33
34
35
36

37 18
38
39 19 **Word Count:** 3,748/4,000
40
41

42 20 **Tables and Figures:** (4/5)
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 27 **ABSTRACT (277/300)**
4

5 28 **Objectives:** Hospitalizations for serious infections are common among middle age and older adults and
6
7 29 frequently used as study outcomes. Yet few studies have evaluated the performance of diagnosis codes to
8
9 30 identify serious infections in this population. We sought to determine the positive predictive value (PPV)
10
11 31 of diagnosis codes for identifying hospitalizations due to serious infections among middle age and older
12
13 32 adults.

14
15 33 **Setting and participants:** We identified hospitalizations for possible infection among adults ≥ 50 years
16
17 34 enrolled in the Tennessee Medicaid healthcare program (2008-2012) using ICD-9 diagnosis codes for
18
19 35 pneumonia, meningitis/encephalitis, bacteremia/sepsis, cellulitis/soft-tissue infections, endocarditis,
20
21 36 pyelonephritis and septic arthritis/osteomyelitis.

22
23 37 **Design:** Medical records were systematically obtained from hospitals randomly selected from a stratified
24
25 38 sampling framework based on geographical region and hospital discharge volume.

26
27 39 **Measures:** Two trained clinical reviewers used a standardized extraction form to abstract information
28
29 40 from medical records. Pre-defined algorithms served as reference to adjudicate confirmed infection-
30
31 41 specific hospitalizations. We calculated the PPV of diagnosis codes using confirmed hospitalizations as
32
33 42 reference. Sensitivity analyses determined the robustness of the PPV to definitions that required
34
35 43 radiological or microbiological confirmation. We also determined interrater reliability between reviewers.

36
37 44 **Results:** The PPV of diagnosis codes for hospitalizations for infection (n=716) was 90% (95% CI: 88-
38
39 45 92). The PPV was highest for pneumonia [97% (95% CI: 94-98)] and cellulitis [91% (95% CI: 85-95)],
40
41 46 and lowest for meningitis/encephalitis [50% (95% CI: 24-76)]. The adjudication reliability was excellent
42
43 47 [93% agreement; first agreement-coefficient: 0.91]. The overall PPV was lower when requiring
44
45 48 microbiological confirmation [45%] and when requiring radiological confirmation for pneumonia [79%].

46
47 49 **Conclusions:** Discharge diagnosis codes have a high PPV for identifying hospitalizations for common,
48
49 50 serious infections among middle age and older adults. PPV estimates for rare infections were imprecise.

50
51 51

52
53 52

53 STRENGTHS AND LIMITATIONS OF THE STUDY

- 54 • This study examined the performance of diagnosis coding algorithms to identify hospitalizations
55 due to serious infections among middle age and older adults enrolled in a State Medicaid program
56 using a systematic and representative sample of records from hospitals of different sizes and in
57 distinct State regions.
- 58 • The reference criteria to identify true infections was based on previous literature and clinical
59 expertise but may be imperfect. Nevertheless, identifying microbiologically-confirmed infections
60 is difficult due to the low sensitivity of culture-based diagnostic methods often used in clinical
61 practice.
- 62 • Diagnosis codes were based on the ICD-9-coding system only. These findings will continue to be
63 helpful for retrospective studies that encompass periods of ICD-9 use, yet additional studies
64 evaluating the performance of ICD-10-based codes would be beneficial.
- 65 • Our coding algorithms to identify serious infections had a high positive predictive value overall,
66 and will be useful in ongoing and future research using administrative data

78 INTRODUCTION

79 Infectious diseases remain a leading cause of morbidity and mortality in the U.S. and
80 elsewhere.(1) Middle age and older adults, in particular, are at high risk for serious infections and their
81 long-term consequences.(2, 3) Among older adults, community-acquired serious infections (including
82 pneumonia, sepsis, and meningitis) often require hospitalization and represent a substantial burden on the
83 U.S. healthcare system.(4-7) The incidence of community-acquired pneumonia is very high among adults
84 ≥ 50 years of age (248 cases per 100,000 adults) with an even higher burden among adults ≥ 80 years of
85 age (1,643 cases per 100,000 adults).(8) Sepsis, cellulitis and pyelonephritis are also very common
86 (sepsis: 100 cases per 100,000 and cellulitis/pyelonephritis: >150 hospitalizations per 100,000 adults)
87 with an increasing incidence of severe sepsis with increased age.(9-11) Meningitis and endocarditis are
88 relatively rare (around 2-3 cases per 100,000), although the case fatality rate is very high.(12, 13)
89 Therefore, it is important to monitor the incidence of these infections, identify important risk factors, and
90 determine the impact of preventative policies (e.g., vaccination) on these diseases among middle age and
91 older adults.(14-16)

92 Large-scale epidemiological studies using administrative data often use serious infections as
93 outcomes.(17-21) However, few studies have evaluated the performance of diagnosis codes to identify
94 serious infections among middle age and older adults. Most previous studies that have assessed the
95 performance of coded discharge diagnosis codes to identify serious infections have focused mainly on
96 common infections (e.g., pneumonia or sepsis), specific populations (e.g., patients with rheumatoid
97 arthritis), or on healthcare-associated or hospital-acquired infections.(22-31) Nevertheless, the
98 performance of coded discharge diagnoses for accurately identifying infections requiring hospitalization
99 among middle age and older adults is unclear. Therefore, we sought to determine the positive predictive
100 value (PPV) of specific discharge diagnoses for identifying infections that required hospitalization among
101 middle age and older adults.

102 103 METHODS

104 **Data sources**

105 TennCare is the managed Medicaid program in the State of Tennessee that provides healthcare
106 insurance to those who are Medicaid eligible (around 20% of the Tennessee population).(32) The adult
107 TennCare population consists of low-income pregnant women and individuals who are elderly or have a
108 disability (over 600,000 annually).(32) We used data from TennCare, supplemented with data from the
109 Tennessee Hospital Discharge Data System (a registry for all hospitalizations in Tennessee) and
110 pharmacy information from Medicare Part D for those that were dual eligible, to identify a retrospective
111 cohort of TennCare enrollees ≥ 50 years of age with pharmacy benefits (2008-2012). We restricted the
112 hospitalizations for serious infection to those occurring from 2008 through 2012 to only include more
113 recent hospitalizations for which medical records are more likely to be available. Cohort members had at
114 least 180 days of baseline continuous enrollment before cohort entry, and were also required to be free of
115 certain life-threatening conditions known to increase the risk of infection (solid organ transplantation,
116 end-stage renal disease, HIV/AIDS, malignancy and serious kidney, liver and respiratory disease) that
117 may limit longitudinal follow-up and impact the assessments of patients' exposures and their risk of
118 infections. Cohort members were also required to have evidence of at least one pharmacy prescription fill
119 and evidence of at least one healthcare encounter during baseline (to ensure use of benefits so that if a
120 healthcare encounter for an infection occurred, it would be detected). Follow-up started on the earliest
121 date the inclusion criteria were met and continued through the earliest of the following: study end date
122 (December 31, 2012), the day prior to diagnosis of a serious life-threatening condition that would have
123 precluded entry to the cohort, loss of enrollment, or date of death. From this retrospective cohort, we
124 identified possible hospitalizations for serious infections (see *Identification of hospitalizations for serious*
125 *infection*) for our validation study. To avoid including infections that may have originated due to a
126 previous hospital stay, we excluded hospitalizations for infections that occurred in the 30-day period after
127 discharge from a previous hospitalization. The study was approved by the institutional review boards of
128 Vanderbilt University and the Tennessee Department of Health, and by the Division of TennCare.

129 **Identification of hospitalizations for serious infection**

1
2
3 130 Clinical knowledge and a literature review were used to identify primary discharge diagnosis
4
5 131 codes that have been used previously to identify specific serious infections that require hospitalization
6
7 132 (*study infections*), including pneumonia (alone or with a primary diagnosis of bacteremia/sepsis),
8
9 133 bacteremia/sepsis, pyelonephritis, meningitis/encephalitis, osteomyelitis/septic arthritis, endocarditis and
10
11 134 cellulitis.(31, 33-35) Specific International Classifications of Diseases-Clinical Modification, 9th-revision
12
13 135 (ICD-9-CM) diagnosis codes used to identify possible hospitalizations for each infection type are
14
15 136 presented in Table 1. As the objective of our study was to determine the PPV of coding algorithms to
16
17 137 identify serious infections that required hospitalization, we focused only on primary diagnoses of
18
19 138 infection to reduce the possibility of detecting concurrent infections that may not have led to
20
21 139 hospitalization or nosocomial infections that developed during the course of the hospitalization.(35)

24 140 **Sampling Strategy**

26 141 We used stratified random sampling to select a representative subset of study infection
27
28 142 hospitalizations from among all possible cases identified in the retrospective cohort from among hospitals
29
30 143 within 200 miles of Vanderbilt University Medical Center (VUMC). Since larger hospitals would be
31
32 144 over-represented in a purely random sampling, and because there may also be regional variability in
33
34 145 coding practices and infection prevalence, we constructed a sampling framework where hospitals were
35
36 146 stratified based on their geographic region in Tennessee (West, Central, and East), and tertiles of reported
37
38 147 discharge volume (Low, Medium, and High) during the study period.(36-38) From this sampling
39
40 148 framework, we randomly selected three hospitals from each of these nine sampling strata, and retrieved
41
42 149 their medical records for review and validation (Figure 1). This strategy, relative to a purely random
43
44 150 sample, ensured better representation of infections identified in smaller hospitals and those in more rural
45
46 151 regions of the State of Tennessee. If a hospital refused to participate, it was replaced by another hospital
47
48 152 randomly selected from the same sampling stratum.

51 153 The overall goal was to review and validate 675 hospitalizations for serious infection from 27
52
53 154 hospitals (25 hospitalizations for each of the 3 hospitals comprising a stratum, yielding 75 hospitalizations
54
55 155 for each of the 9 strata) (Figure 1). We conservatively assumed that up to 80% of records requested would

1
2
3 156 be available for review, and so we requested 32 records per hospital to receive an average of 25 records
4
5 157 from each (Figure 1). To ensure that we reviewed sufficient rare infections, we preferentially selected any
6
7 158 identified possible hospitalizations for meningitis/encephalitis, osteomyelitis/septic arthritis and
8
9 159 endocarditis from each hospital in the sample. We randomly selected the remaining set of possible
10
11 160 hospitalizations for other serious infections based on the proportional distribution of common infections
12
13 161 at each hospital (pneumonia, bacteremia/sepsis, pyelonephritis and cellulitis) until 32 infections were
14
15 162 identified. For hospitals with fewer than 32 infections during the study period, all infections were
16
17 163 requested.

18 164 **Abstraction of Medical Records**

19
20 165 Relevant clinical information was abstracted from the medical record (transfer notes, emergency
21
22 166 room summary, admission summary, physical/history, pharmacy information, laboratory, microbiology,
23
24 167 and radiology information, and discharge summary) of each hospitalization with a primary discharge
25
26 168 diagnosis code indicative of infection using a standardized and customized REDCap electronic data
27
28 169 capture instrument hosted at Vanderbilt University.⁽³⁹⁾ As we were interested in infections that led to
29
30 170 hospitalizations, we focused our reviews on clinical, microbiological and radiological information from
31
32 171 the 2 days prior to the admission date through 2 days after admission to limit the possibility of identifying
33
34 172 infections that developed during the hospitalization (i.e. nosocomial infections). In preparation for this
35
36 173 study, the case report form was pilot-tested among a separate, convenience sample of 354 possible
37
38 174 infections identified in the cohort from 3 hospitals in the same city as Vanderbilt University. This separate
39
40 175 sample of hospitalizations was used only for pilot-testing the case report form, and was not included in
41
42 176 the current study. One trained medical reviewer abstracted the relevant information for all selected
43
44 177 records using the case report form. During the abstraction process, the lack of a particular finding in the
45
46 178 medical record was treated as a lack of evidence for that finding, and so no information was considered
47
48 179 missing after abstraction.

49 180 **Adjudication of Medical Records**

1
2
3 181 All records received were abstracted, reviewed and adjudicated. We made the final determination
4
5 182 of whether a hospitalization represented a confirmed infection or not using *a priori* definitions of clinical,
6
7 183 radiological, and/or microbiological findings compatible with infection for each infection type. Previous
8
9 184 validation studies and expert clinical knowledge were used to define these specific *a priori* definitions for
10
11 185 each infection type (*Supplementary appendix*). (31, 35, 40)

13 186 **Statistical analysis**

15
16 187 We calculated the PPV of the ICD-9-CM discharge diagnosis codes for identifying
17
18 188 hospitalizations for serious infection using the results of the *a priori* definitions applied to the information
19
20 189 abstracted from the medical records as the reference (the proportion of cases identified with discharge
21
22 190 diagnosis codes that were determined to be true cases after adjudication of the medical record
23
24 191 information). We calculated 95% confidence intervals for the PPV using Wilson's formula. (41)
25
26 192 Secondary analyses assessed the PPV for hospitalizations for serious infection across hospitals of
27
28 193 different sizes and in different geographical regions of Tennessee.

29
30 194 We also assessed the reliability of the abstraction process. A second trained medical reviewer
31
32 195 abstracted relevant information from a subset of selected records, which included all meningitis and
33
34 196 endocarditis records, and a random selection of 10% of each of the remaining infection types. Each
35
36 197 reviewer conducted the process independently and blinded from one another. For the subset of records
37
38 198 abstracted by both reviewers, inter-reviewer agreement for the adjudication of a true or mis-identified
39
40 199 infection was assessed using the Gwet's first agreement coefficient (AC_1). (42-44) Since Cohen's kappa
41
42 200 statistic can be unreliable when the prevalence of the event and the level of observer agreement are high
43
44 201 in the study sample, we used Gwet's AC_1 as a reliability measure unlikely to be affected by these
45
46 202 concerns. (44-46)

47
48
49 203 In planned sensitivity analyses, we first assessed the impact of excluding hospitalizations that
50
51 204 occurred after the individual was transferred from another healthcare facility, as initial documentation and
52
53 205 details of the infection could be missing or incomplete in the receiving hospital. (40) We also assessed the
54
55 206 impact on the PPV for all infections when requiring microbiological identification of a pathogen

207 (excluding common contaminants) from a sterile site within 2 days before or after the hospitalization
208 admission date. A final sensitivity analysis among hospitalizations for possible pneumonia assessed the
209 PPV when radiological evidence of pneumonia was required [i.e. pneumonia, opacity, or infiltrate
210 mentioned in a chest X-ray or computed tomography scan report] (*Supplementary appendix*). All analyses
211 were performed in Stata-IC, version 15.1 (College Station TX).

212 Patient and Public Involvement

213 No patients were involved in the development of the research question, the outcome measures, or the
214 design or conduct of the study. As we conducted a retrospective study using administrative data, we have
215 no plans to disseminate the results of the research to study participants.

217 RESULTS

218 *Cohort characteristics*

219 Among a retrospective cohort of 129,465 adults ≥ 50 years of age enrolled in TennCare, 9,769
220 hospitalizations for serious infection were identified during the study period (2008-2012) among 7,770
221 unique patients (Figure 2). Cohort members were primarily female (57.8%) with a median age of 54 years
222 (mean: 57 years; range: 50-110). Among the 8,322 hospitalizations for serious infection that occurred at a
223 hospital within 200 miles of VUMC, pneumonia, cellulitis and bacteremia/sepsis were the most common
224 infections (54.3%, 20.5% and 18.4%, respectively), followed by pyelonephritis (3.8%) and septic
225 arthritis/osteomyelitis (2.5%). Fewer than 1% of hospitalizations were due to meningitis/encephalitis
226 (n=30) and endocarditis (n=18).

227 *Collection, review and adjudication of selected medical records*

228 Of the 27 hospitals that were initially selected for the sample, 21 (78%) were able to participate.
229 We selected 7 additional hospitals to replace the 6 non-participants to achieve the desired sample size,
230 including an additional small hospital in the East region due to a large number of unavailable records
231 from a single participating hospital.

232 We received 716 (89%) of 808 requested records from 28 participating hospitals [Table 2].
233 Record availability from participating hospitals was lower in medium size hospitals (81.8%) compared to
234 small (93.5%) and large hospitals (91.7%), but did not differ by geographic region. Record availability by
235 infection type was greater than 86% for all infection types, with the exception of hospitalizations for the
236 rare endocarditis cases (57.1%; only 4 of 7 cases).

237 The sample of hospitalizations for serious infection included patients who were primarily female
238 (63.6%), with a median age of 60 years (mean: 64 years; range: 50-101) at the time of hospitalization.
239 There was evidence of transfer from a prior healthcare facility for 21.8% of the hospitalizations for
240 serious infection [highest percentage of transfers for bacteremia/sepsis (38.5%) and pneumonia (25.1%)].
241 The most common healthcare facility source was a nursing home/skilled nursing facility (84.6%), but also
242 included group home sources (7.7%), other sources (4.5%) [assisted living facility, mental health center]
243 and another acute care hospital (3.2%). There was evidence of an emergency department visit within 7
244 days prior to admission date for the serious infection hospitalization in 4.8% of the records.

245 *Performance of discharge diagnosis codes*

246 A total of 646 [PPV: 90.2% (95% CI: 87.8-92.2)] of the hospitalizations for serious infection
247 identified using ICD-9-CM primary discharge diagnosis codes were confirmed by applying the *a priori*
248 definitions to the abstracted data. The PPV was highest for pneumonia and cellulitis [96.5% (95% CI:
249 93.9-98.0) and 91.1% (95% CI: 84.7-94.9), respectively], and was $\geq 75\%$ for bacteremia/sepsis,
250 pyelonephritis, septic arthritis/osteomyelitis, and endocarditis. The PPV was lowest for
251 meningitis/encephalitis [50.0% (95% CI: 23.7-76.3)], although the precision was limited due to a low
252 number of available records for review (Table 2). Among the 10 potential cases of
253 meningitis/encephalitis, 7 cases were meningitis/meningoencephalitis and 3 were encephalitis. The
254 respective PPVs for meningitis/meningoencephalitis and encephalitis were 71.4% (95% CI: 35.9-91.8)
255 and 0%, respectively.

256 When performance was evaluated across stratification sampling parameters, no apparent
257 differences were observed in the PPV for records from hospitals in different geographical regions of

1
2
3 258 Tennessee. Although the PPV was high for all three discharge volume groups, the PPV was significantly
4
5 259 lower in large hospitals [84.6% (95% CI: 79.6-88.5)] compared to smaller hospitals [93.9% (95% CI:
6
7 260 90.0-96.3); PPV difference: -9.3% (95% CI: -14.7, -3.9)] and medium hospitals [92.7% (95% CI: 88.6-
8
9 261 95.4); PPV difference: -8.1% (95% CI: -13.7, -2.6)] (Table 2). This was likely driven by the different
10
11 262 distributions in the types of infections selected for review in the hospital groups. Large hospitals had a
12
13 263 higher proportion of non-pneumonia infections (70.4%) compared to medium and small hospitals (49.4%
14
15 264 and 36.1%, respectively). Importantly, the PPV for pneumonia was similar in each discharge volume
16
17 265 group (range: 96.0 to 96.6%), whereas the PPV was smaller for non-pneumonia infections in large
18
19 266 hospitals (79.8%) compared to medium (88.7%) and small (89.2%) hospitals.

20
21
22 267 In the 82 records independently abstracted by two reviewers to assess reliability, there was 92.7%
23
24 268 (95% CI: 86.9-98.4) agreement for identifying true hospitalizations for serious infection. The inter-rater
25
26 269 agreement was also high when assessing reliability, independent of the outcome prevalence, with an AC₁
27
28 270 of 0.91 (95% CI: 0.84-0.99). Of the 6 discordant cases, 3 were meningitis/encephalitis (1
29
30 271 meningitis/meningoencephalitis and 2 encephalitis), with one each of bacteremia/sepsis, pyelonephritis
31
32 272 and septic arthritis. The main reason for a discrepancy between reviewers was whether or not treatment
33
34 273 for the infection of interest occurred within 2 days of the admission date, which was one of the major
35
36 274 criteria for adjudication (see *Supplementary appendix*).

37 275 *Sensitivity analyses*

38
39
40
41 276 The PPV was virtually unchanged when excluding the 21.8% of hospitalizations that occurred as
42
43 277 transfers from another healthcare facility [90.1% (95% CI: 87.4-92.3)]. Microbiological evidence of the
44
45 278 specific infection type was found in 47.6% of records, leading to reduced PPVs when requiring
46
47 279 microbiological evidence [45.4% (95% CI: 41.8-49.1)]. Microbiological evidence of infection was
48
49 280 highest in hospitalizations for suspected pyelonephritis (94.4%), but was $\leq 60\%$ for every other infection
50
51 281 type [pneumonia (42.7%); cellulitis/soft tissue infections (58.5%); bacteremia/sepsis (26.1%)]. When
52
53 282 requiring radiological confirmation of pneumonia, the PPV for coded diagnoses was 78.8% (95% CI:
54
55 283 74.2-82.8). Approximately 95.6% of possible hospitalizations for pneumonia had at least one documented

284 chest x-ray or CT-scan. Among those patients with a chest x-ray or CT-scan report available (n=325),
285 83.4% had a finding compatible with pneumonia. The main findings among the 54 patients with possible
286 pneumonia and a radiological report available, but without radiological confirmation of pneumonia
287 included atelectasis (n=6), interstitial pneumonitis (n=3), chronic heart failure with pulmonary edema
288 (n=1), and no radiological findings of any kind (n=44).

289

290 DISCUSSION

291 Discharge diagnoses for identifying hospitalizations due to serious infections among middle age
292 and older adults had an overall PPV of 90.2%, with highest values for the identification of common
293 serious infections. PPVs were consistently high across different hospital types and regions of Tennessee.
294 Furthermore, the PPV was similar after exclusion of hospitalizations for serious infection that were the
295 result of a transfer from another healthcare facility (e.g. acute care hospital, skilled nursing facility).
296 Microbiological confirmation was available for fewer than 50% of patients admitted with possible
297 hospitalizations for serious infections, and as expected, the requirement resulted in a low PPV for all
298 infections, with the exception of pyelonephritis. Importantly, the PPV for pneumonia hospitalizations
299 remained relatively high even when requiring radiological confirmation.

300 The PPV for hospitalizations for pneumonia in previous smaller validation studies has ranged
301 from 72 to 86% in different healthcare systems, but those studies were not focused on middle age and
302 older adults.(31, 47-49) In our study of hospitalizations among middle age and older adults, we found that
303 coded discharge diagnoses have a higher PPV for pneumonia compared to previous studies. The PPV for
304 bacteremia/sepsis was also on the higher range of previously reported PPVs for diagnosis codes to
305 identify bacteremia/sepsis from administrative data in other populations (reported range from 45% to
306 97.7%), and for septic arthritis/osteomyelitis compared to a previous study conducted among patients
307 with diabetes (63.9% versus 75.9% in our study).(23, 50, 51) Overall, the observed PPV for all infections
308 in our study was comparable to two previous comprehensive validation studies of bacterial infections, one
309 among patients with rheumatoid arthritis in a single hospital system and another among patients in one of

12

1
2
3 310 the Veteran's Affairs integrated service networks.(35, 40) Compared to these two previous studies of
4
5 311 ICD-9 codes, we abstracted and adjudicated a larger number of records while using a more systematic
6
7 312 sampling strategy to retrieve and review records for hospitalizations from multiple regions and hospital
8
9 313 types as opposed to a single hospital or healthcare system. However, some of the PPVs for individual
10
11 314 infections were less precise and less similar to these previous studies. This was especially true for rare
12
13 315 infections, as would be expected due to the low numbers of rare infections in our study and across
14
15 316 previous studies.(35, 40) The results of our study are also similar to previous validation studies that used
16
17 317 corresponding ICD-10 diagnosis codes to identify hospitalizations for serious infection.(52, 53)

18
19
20 318 One limitation to consider in our study was that it was not designed to estimate the sensitivity and
21
22 319 specificity of the coding algorithms. This would have required the identification, review and adjudication
23
24 320 of a sample of hospitalizations that did not fulfill our algorithm (i.e. presence of the ICD-9 primary
25
26 321 discharge diagnosis codes indicative of infection). However, when the prevalence of an outcome is low,
27
28 322 the PPV approximates the specificity.(54) Importantly, any non-differential outcome misclassification
29
30 323 between exposure groups resulting from the use of imperfect but highly-specific measurements would
31
32 324 attenuate the impact of the misclassification on the relative risk estimates.(55) In addition, we found that
33
34 325 the PPV of coded discharge diagnoses for serious infections remained high across hospitals of different
35
36 326 sizes and across different geographical areas of Tennessee, which may have different rates of
37
38 327 hospitalizations for serious infection.(56) Although our study applied a systematic sampling strategy to
39
40 328 assure the representation of different settings in our population, our population was restricted to middle
41
42 329 age and older adults enrolled in a State Medicaid program. Therefore, caution is warranted when
43
44 330 extrapolating the study findings to other populations.

45
46
47 331 Another limitation is the use of available clinical information to operationalize definitions for
48
49 332 adjudication of true hospitalizations for infections. It is possible that some procedures, laboratory findings
50
51 333 and diagnoses that informed the final diagnosis of infection were not fully recorded in the medical
52
53 334 records, and thus, were not available for our review and may have contributed to the observed PPV for
54
55 335 some infections. Although we used previous validation studies and clinical information to build pre-

1
2
3 336 specified definitions for the adjudication of true infections, our reference criteria may be imperfect,
4
5 337 considering the retrospective nature of our determinations and potential variability in clinical practice.
6
7 338 Nevertheless, we also assessed how the availability of selected findings (i.e. microbiological and
8
9 339 radiological information) in the medical record impacted the overall and infection-specific PPV. We
10
11 340 demonstrated that relying on highly specific clinical diagnostics, such as microbiological and radiological
12
13 341 information, to confirm true infections would result in lower PPVs for identification of infections in
14
15 342 administrative data. Requiring microbiological confirmation to confirm true infections is challenging
16
17 343 because of the known low sensitivity of culture-based diagnostic methods (most commonly used in
18
19 344 clinical practice), which may lead to misclassification.^(57, 58) In addition, requiring radiological evidence
20
21 345 compatible with pneumonia within 2 days of hospital admission did lower the observed PPV for
22
23 346 pneumonia hospitalizations. Nevertheless, the observed PPV remained close to 80%, which should reduce
24
25 347 concerns about using diagnosis codes to identify hospitalizations due to pneumonia. Finally, the coding
26
27 348 algorithms were based on the ICD-9-coding system only. Although these findings will be helpful for
28
29 349 retrospective studies that encompass periods of ICD-9 use, additional studies evaluating the performance
30
31 350 of ICD-10-based codes would be useful to complement our findings.
32
33

34
35 351 Our study demonstrated that discharge diagnosis codes can be used to accurately identify
36
37 352 hospitalizations for serious infections among middle age and older adults. The highest PPVs were
38
39 353 observed for the most common infections, and the PPV for pneumonia remained high when requiring
40
41 354 radiological confirmation. Importantly, consistently high PPVs were observed across different hospital
42
43 355 sizes and regions. However, the estimated PPV was lower and less precise for very rare infections (e.g.
44
45 356 encephalitis). This should be an important consideration for studies specifically focused on those less
46
47 357 frequent outcomes, especially when strict microbiological confirmation is required. Taken together, these
48
49 358 findings support the use of discharge diagnosis codes for infections to identify outcomes in ongoing and
50
51 359 future epidemiological studies among middle age and older adults.
52

53
54 360

55 361 **Acknowledgement**

56
57
58 14
59

1
2
3 362 We are indebted to the Tennessee Bureau of TennCare of the Tennessee Department of Finance and
4
5 363 Administration, which provided data for the study. We are also indebted to the Tennessee Department of
6
7 364 Health for providing data for the study. Statements in the report should not be construed as endorsement
8
9 365 by the U.S. Department of Health and Human Services, the Department of Veterans Affairs, or the
10
11 366 Tennessee Department of Health.
12

13 367

16 368 **Funding**

17
18 369 This study was funded by the NIH (R03-AG-042981 and R01-AG-043471-01A1) and the TL1 award
19
20 370 TL1TR000447.
21

22 371 **Contributors**

23
24 372 ADW planned the medical record collection and statistical analysis, analyzed and interpreted the data, and
25
26 373 drafted and revised the paper. MRG, WS, CMS, and RAG planned the statistical analysis, interpreted the
27
28 374 data and revised the paper. EFM prepared the data, and revised the paper. CGG designed the project,
29
30 375 acquired the data from TennCare, planned the medical record collection and statistical analysis,
31
32 376 interpreted the data, and revised the paper.
33

34 377 **Declaration of interests**

35
36 378 CGG has received consulting fees from Pfizer and Merck, and received research support from Sanofi-
37
38 379 Pasteur, Campbell Alliance, the Centers for Disease Control and Prevention, National Institutes of Health,
39
40 380 The Food and Drug Administration, and the Agency for Health Care Research and Quality. WS has
41
42 381 received personal fees from Pfizer, Merck, Novavax, Dynavax, Sanofi-Pasteur, GSK, Seqirus, and
43
44 382 received research support from the Centers for Disease Control and Prevention. ADW, EFM, CMS, MRG
45
46 383 and RAG have no conflicts of interest to disclose.
47
48

49 384 **Data sharing**

50
51 385 No additional unpublished data are available from the study. The study protocol and statistical code are
52
53 386 available from the corresponding author, Andrew Wiese (andrew.d.wiese@vanderbilt.edu).
54
55

56 387

388 **REFERENCES**

- 389 1. Khabbaz RF, Moseley RR, Steiner RJ, Levitt AM, Bell BP. Challenges of infectious diseases in
390 the USA. *Lancet*. 2014;384(9937):53-63.
- 391 2. Grohskopf LA, Sokolow LZ, Broder KR, Olsen SJ, Karron RA, Jernigan DB, et al. Prevention
392 and Control of Seasonal Influenza with Vaccines. *MMWR Recomm Rep*. 2016;65(5):1-54.
- 393 3. Kim DK, Bridges CB, Harriman KH. Advisory committee on immunization practices
394 recommended immunization schedule for adults aged 19 years or older--United States, 2015.
395 *MMWR Morb Mortal Wkly Rep*. 2015;64(4):91-2.
- 396 4. Jeon MH, Chung JW, Choi SH, Kim TH, Lee EJ, Choo EJ. Pneumonia risk factors and clinical
397 features of hospitalized patients older than 15 years with pandemic influenza A (H1N1) in South
398 Korea: a multicenter study. *Diagn Microbiol Infect Dis*. 2011;70(2):230-5.
- 399 5. Pilishvili T, Bennett NM. Pneumococcal Disease Prevention Among Adults: Strategies for the
400 Use of Pneumococcal Vaccines. *Am J Prev Med*. 2015;49(6 Suppl 4):S383-90.
- 401 6. Sjoding MW, Prescott HC, Wunsch H, Iwashyna TJ, Cooke CR. Longitudinal Changes in ICU
402 Admissions Among Elderly Patients in the United States. *Crit Care Med*. 2016;44(7):1353-60.
- 403 7. Crotty MP, Meyers S, Hampton N, Bledsoe S, Ritchie DJ, Buller RS, et al. Epidemiology, Co-
404 Infections, and Outcomes of Viral Pneumonia in Adults: An Observational Cohort Study.
405 *Medicine (Baltimore)*. 2015;94(50):e2332.
- 406 8. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM, et al. Community-Acquired
407 Pneumonia Requiring Hospitalization among U.S. Adults. *N Engl J Med*. 2015;373(5):415-27.
- 408 9. Filbin MR, Arias SA, Camargo CA, Jr., Barche A, Pallin DJ. Sepsis visits and antibiotic
409 utilization in U.S. emergency departments*. *Crit Care Med*. 2014;42(3):528-35.
- 410 10. Christensen KL, Holman RC, Steiner CA, Sejvar JJ, Stoll BJ, Schonberger LB. Infectious disease
411 hospitalizations in the United States. *Clin Infect Dis*. 2009;49(7):1025-35.

- 1
2
3 412 11. Ginde AA, Moss M, Shapiro NI, Schwartz RS. Impact of older age and nursing home residence
4
5 413 on clinical outcomes of US emergency department visits for severe sepsis. *J Crit Care.*
6
7 414 2013;28(5):606-11.
8
9 415 12. Lung B, Vahanian A. Epidemiology of acquired valvular heart disease. *Can J Cardiol.*
10
11 416 2014;30(9):962-70.
12
13 417 13. Thigpen MC, Whitney CG, Messonnier NE, Zell ER, Lynfield R, Hadler JL, et al. Bacterial
14
15 418 meningitis in the United States, 1998-2007. *N Engl J Med.* 2011;364(21):2016-25.
16
17 419 14. Simonetti AF, Garcia-Vidal C, Viasus D, Garcia-Somoza D, Dorca J, Gudiol F, et al. Declining
18
19 420 mortality among hospitalized patients with community-acquired pneumonia. *Clin Microbiol*
20
21 421 *Infect.* 2016;22(6):567.e1-7.
22
23 422 15. Rivero-Calle I, Pardo-Seco J, Aldaz P, Vargas DA, Mascaros E, Redondo E, et al. Incidence and
24
25 423 risk factor prevalence of community-acquired pneumonia in adults in primary care in Spain
26
27 424 (NEUMO-ES-RISK project). *BMC Infect Dis.* 2016;16(1):645.
28
29 425 16. Park JY, Park S, Lee SH, Lee MG, Park YB, Oh KC, et al. Microorganisms Causing Community-
30
31 426 Acquired Acute Bronchitis: The Role of Bacterial Infection. *PLoS One.* 2016;11(10):e0165553.
32
33 427 17. Jackson ML, Walker R, Lee S, Larson E, Dublin S. Predicting 2-Year Risk of Developing
34
35 428 Pneumonia in Older Adults without Dementia. *J Am Geriatr Soc.* 2016;64(7):1439-47.
36
37 429 18. Clement RC, Haddix KP, Creighton RA, Spang JT, Tennant JN, Kamath GV. Risk Factors for
38
39 430 Infection After Knee Arthroscopy: Analysis of 595,083 Cases From 3 United States Databases.
40
41 431 Arthroscopy. 2016;32(12):2556-61.
42
43 432 19. Weinert BA, Edmonson MB. Hospitalizations at Nonfederal Facilities for Lower Respiratory
44
45 433 Tract Infection in American Indian and Alaska Native Children Younger than 5 Years of Age,
46
47 434 1997-2012. *J Pediatr.* 2016;175:33-9.e4.
48
49 435 20. Akiyama T, Chikuda H, Yasunaga H, Horiguchi H, Fushimi K, Saita K. Incidence and risk
50
51 436 factors for mortality of vertebral osteomyelitis: a retrospective analysis using the Japanese
52
53 437 diagnosis procedure combination database. *BMJ Open.* 2013;3(3).

- 1
2
3 438 21. Graversen ME, Dalgaard LS, Jensen-Fangel S, Jespersen B, Ostergaard L, Sogaard OS. Risk and
4
5 439 outcome of pyelonephritis among renal transplant recipients. *BMC Infect Dis.* 2016;16:264.
6
7 440 22. Abou Zahr Z, Spiegelman A, Cantu M, Ng B. Perioperative use of anti-rheumatic agents does not
8
9 441 increase early postoperative infection risks: a Veteran Affairs' administrative database study.
10
11 442 *Rheumatol Int.* 2015;35(2):265-72.
12
13 443 23. Carnahan RM, Herman RA, Moores KG. A systematic review of validated methods for
14
15 444 identifying transfusion-related sepsis using administrative and claims data. *Pharmacoepidemiol*
16
17 445 *Drug Saf.* 2012;21 Suppl 1:222-9.
18
19 446 24. Condell O, Gubbels S, Nielsen J, Espenhain L, Frimodt-Moller N, Engberg J, et al. Automated
20
21 447 surveillance system for hospital-acquired urinary tract infections in Denmark. *J Hosp Infect.*
22
23 448 2016;93(3):290-6.
24
25 449 25. Gedeborg R, Furebring M, Michaelsson K. Diagnosis-dependent misclassification of infections
26
27 450 using administrative data variably affected incidence and mortality estimates in ICU patients. *J*
28
29 451 *Clin Epidemiol.* 2007;60(2):155-62.
30
31 452 26. Goto M, Ohl ME, Schweizer ML, Perencevich EN. Accuracy of administrative code data for the
32
33 453 surveillance of healthcare-associated infections: a systematic review and meta-analysis. *Clin*
34
35 454 *Infect Dis.* 2014;58(5):688-96.
36
37 455 27. Shaklee J, Zerr DM, Elward A, Newland J, Leckerman K, Asti L, et al. Improving surveillance
38
39 456 for pediatric *Clostridium difficile* infection: derivation and validation of an accurate case-finding
40
41 457 tool. *Pediatr Infect Dis J.* 2011;30(3):e38-40.
42
43 458 28. Wright SB, Huskins WC, Dokholyan RS, Goldmann DA, Platt R. Administrative databases
44
45 459 provide inaccurate data for surveillance of long-term central venous catheter-associated
46
47 460 infections. *Infect Control Hosp Epidemiol.* 2003;24(12):946-9.
48
49 461 29. Greenberg JA, Hohmann SF, Hall JB, Kress JP, David MZ. Validation of a Method to Identify
50
51 462 Immunocompromised Patients with Severe Sepsis in Administrative Databases. *Ann Am Thorac*
52
53 463 *Soc.* 2016;13(2):253-8.

- 1
2
3 464 30. Olsen MA, Young-Xu Y, Stwalley D, Kelly CP, Gerding DN, Saeed MJ, et al. The burden of
4
5 465 clostridium difficile infection: estimates of the incidence of CDI from U.S. Administrative
6
7 466 databases. *BMC Infect Dis.* 2016;16:177.
8
9 467 31. Grijalva CG, Chung CP, Stein CM, Gideon PS, Dyer SM, Mitchel EF, Jr., et al. Computerized
10
11 468 definitions showed high positive predictive values for identifying hospitalizations for congestive
12
13 469 heart failure and selected infections in Medicaid enrollees with rheumatoid arthritis.
14
15 470 *Pharmacoepidemiol Drug Saf.* 2008;17(9):890-5.
16
17 471 32. TennCare Fiscal Year 2015-2016 Annual Report. In: Administration DoHCF, ed; 2016.
18
19 472 33. Wiese AD, Griffin MR, Stein CM, Mitchel EF, Jr., Grijalva CG. Opioid Analgesics and the Risk
20
21 473 of Serious Infections Among Patients With Rheumatoid Arthritis: A Self-Controlled Case Series
22
23 474 Study. *Arthritis Rheumatol.* 2016;68(2):323-31.
24
25 475 34. Grijalva CG, Kaltenbach L, Arbogast PG, Mitchel EF, Jr., Griffin MR. Initiation of rheumatoid
26
27 476 arthritis treatments and the risk of serious infections. *Rheumatology (Oxford).* 2010;49(1):82-90.
28
29 477 35. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital
30
31 478 discharge databases coded serious bacterial infections accurately. *J Clin Epidemiol.*
32
33 479 2007;60(4):397-409.
34
35 480 36. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values
36
37 481 with disease prevalence. *Stat Med.* 1997;16(9):981-91.
38
39 482 37. Kim HM, Smith EG, Stano CM, Ganoczy D, Zivin K, Walters H, et al. Validation of key
40
41 483 behaviourally based mental health diagnoses in administrative data: suicide attempt, alcohol
42
43 484 abuse, illicit drug abuse and tobacco use. *BMC Health Serv Res.* 2012;12:18.
44
45 485 38. El-Ghitany EM, Farghaly AG, Farag S, Abd El-Wahab EW. Validation of EGCRISC for Chronic
46
47 486 Hepatitis C Infection Screening and Risk Assessment in the Egyptian Population. *PLoS One.*
48
49 487 2016;11(12):e0168649.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 488 39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data
4 489 capture (REDCap)--a metadata-driven methodology and workflow process for providing
5 490 translational research informatics support. *J Biomed Inform.* 2009;42(2):377-81.
6
7 491 40. Patkar NM, Curtis JR, Teng GG, Allison JJ, Saag M, Martin C, et al. Administrative codes
8 492 combined with medical records based criteria accurately identified bacterial infections among
9 493 rheumatoid arthritis patients. *J Clin Epidemiol.* 2009;62(3):321-7, 7.e1-7.
10
11 494 41. Niesner K, Murff HJ, Griffin MR, Wasserman B, Greevy R, Grijalva CG, et al. Validation of VA
12 495 administrative data algorithms for identifying cardiovascular disease hospitalization.
13 496 *Epidemiology.* 2013;24(2):334-5.
14
15 497 42. Govatsmark RE, Sneeggen S, Karlsaune H, Slordahl SA, Bonna KH. Interrater reliability of a
16 498 national acute myocardial infarction register. *Clin Epidemiol.* 2016;8:305-12.
17
18 499 43. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and
19 500 Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with
20 501 personality disorder samples. *BMC Med Res Methodol.* 2013;13:61.
21
22 502 44. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br*
23 503 *J Math Stat Psychol.* 2008;61(Pt 1):29-48.
24
25 504 45. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: resolution of the two
26 505 paradoxes. *J Clin Epidemiol.* 1996;49(4):431-4.
27
28 506 46. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J*
29 507 *Clin Epidemiol.* 1990;43(6):543-9.
30
31 508 47. Drahos J, Vanwormer JJ, Greenlee RT, Landgren O, Koshiol J. Accuracy of ICD-9-CM codes in
32 509 identifying infections of pneumonia and herpes simplex virus in administrative data. *Ann*
33 510 *Epidemiol.* 2013;23(5):291-3.
34
35 511 48. Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients
36 512 with pneumonia. *Am J Med Qual.* 2005;20(6):319-28.

- 1
2
3 513 49. Kern DM, Davis J, Williams SA, Tunceli O, Wu B, Hollis S, et al. Validation of an
4
5 514 administrative claims-based diagnostic code for pneumonia in a US-based commercially insured
6
7 515 COPD population. *Int J Chron Obstruct Pulmon Dis*. 2015;10:1417-25.
8
9 516 50. Iwashyna TJ, Odden A, Rohde J, Bonham C, Kuhn L, Malani P, et al. Identifying patients with
10
11 517 severe sepsis using administrative claims: patient-level validation of the angus implementation of
12
13 518 the international consensus conference definition of severe sepsis. *Med Care*. 2014;52(6):e39-43.
14
15 519 51. Newton KM, Wagner EH, Ramsey SD, McCulloch D, Evans R, Sandhu N, et al. The use of
16
17 520 automated data to identify complications and comorbidities of diabetes: a validation study. *J Clin*
18
19 521 *Epidemiol*. 1999;52(3):199-207.
20
21 522 52. Sahli L, Lapeyre-Mestre M, Derumeaux H, Moulis G. Positive predictive values of selected
22
23 523 hospital discharge diagnoses to identify infections responsible for hospitalization in the French
24
25 524 national hospital database. *Pharmacoepidemiol Drug Saf*. 2016;25(7):785-9.
26
27 525 53. Holland-Bill L, Xu H, Sorensen HT, Acquavella J, Svaerke C, Gammelager H, et al. Positive
28
29 526 predictive value of primary inpatient discharge diagnoses of infection among cancer patients in
30
31 527 the Danish National Registry of Patients. *Ann Epidemiol*. 2014;24(8):593-7, 7.e1-18.
32
33 528 54. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic
34
35 529 research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-37.
36
37 530 55. Rodgers A, MacMahon S. Systematic underestimation of treatment effects as a result of
38
39 531 diagnostic test inaccuracy: implications for the interpretation and design of thromboprophylaxis
40
41 532 trials. *Thromb Haemost*. 1995;73(2):167-71.
42
43 533 56. van Walraven C, English S, Austin PC. Administrative database code accuracy did not vary
44
45 534 notably with changes in disease prevalence. *J Clin Epidemiol*. 2016;79:86-9.
46
47 535 57. Mancini N, Carletti S, Ghidoli N, Cichero P, Burioni R, Clementi M. The era of molecular and
48
49 536 other non-culture-based methods in diagnosis of sepsis. *Clin Microbiol Rev*. 2010;23(1):235-51.
50
51 537 58. Niederman MS. The argument against using quantitative cultures in clinical trials and for the
52
53 538 management of ventilator-associated pneumonia. *Clin Infect Dis*. 2010;51 Suppl 1:S93-9.
54
55
56
57
58 21

539 **Table 1. Discharge diagnosis code definitions (ICD-9-CM) for hospitalizations for serious infection**

Serious Infection	Primary (first listed) discharge diagnosis code
Pneumonia-primary definition	003.22, 480.* [†] , 481, 482.*, 483.*, 484.*, 485.*, 486.*, 487.0
Pneumonia-secondary definition (pneumonia diagnosis (above) in any other diagnosis field)	510.*, 038.*, 790.7, 995.91, 995.92
Meningitis/ Encephalitis	003.21, 036.0, 047.*, 049.*, 053.0, 054.72, 072.1, 091.81, 094.2, 098.82, 100.81, 320.*, 036.1, 054.3, 056.01, 058.21, 058.29, 062.*, 063.*, 064.*, 066.41, 072.2, 094.81, 130.0, 323.*
Bacteremia/ Sepsis [†]	038.*, 790.7, 995.91, 995.92
Cellulitis/ Soft-tissue infections	035, 040.0, 569.61, 681.*, 682.*, 728.86, 785.4
Endocarditis	036.42, 074.22, 093.2*, 098.84, 421.*, 422.92
Pyelonephritis	590.*
Septic Arthritis/ Osteomyelitis	003.23, 056.71, 098.5*, 711.0, 711.00-711.07, 711.09, 711.9*, 003.24, 376.03, 526.4, 730.0*, 730.1*, 730.2*

540 † Without a diagnosis of pneumonia in any other diagnosis field

541 ‡ A * indicates all numeric values [0-9]

542

Table 2. Positive predictive value (PPV) of coded discharge diagnosis definitions for hospitalizations for serious infections among adults ≥ 50 years of age enrolled in Tennessee Medicaid, 2008-2012

Type	Expected Number of Records	Records Received	PPV (95 % CI)
Overall	675	716	90.2 (87.8, 92.2)
Region Specific			
West	225	195	91.3 (86.5, 94.5)
Central	225	225	88.9 (84.1, 92.4)
East	225	296	90.5 (86.7, 93.4)
Bed volume size specific			
Low	225	230	93.9 (90.0, 96.3)
Medium	225	233	92.7 (88.6, 95.4)
High	225	253	84.6 (79.6, 88.5)
Serious Infection			
Pneumonia	305	340	96.5 (93.9, 98.0)
Cellulitis/Soft-tissue infections	125	123	91.1 (84.7, 94.9)
Pyelonephritis	80	89	87.6 (79.2, 93.0)
Bacteremia/Sepsis	100	92	82.6 (73.6, 89.0)
Septic Arthritis/Osteomyelitis	50	58	75.9 (63.5, 85.0)
Meningitis/Encephalitis	10	10	50.0 (23.7, 76.3)
Endocarditis	5	4	75.0 (30.1, 95.4)

Figure legends

Figure 1. Sampling strategy for identifying potential hospitalizations for serious infection

Figure 2. Identifying a retrospective cohort of patients ≥ 50 years of age without serious/life-threatening conditions, Tennessee Medicaid (2008-2012)

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

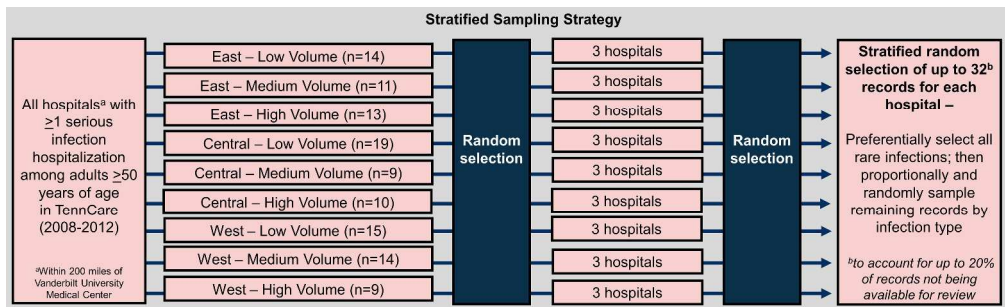


Figure 1. Sampling strategy for identifying potential hospitalizations for serious infection

457x139mm (300 x 300 DPI)

peer review only

BMJ Open: first published as 10.1136/bmjopen-2017-020857 on 19 June 2018. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

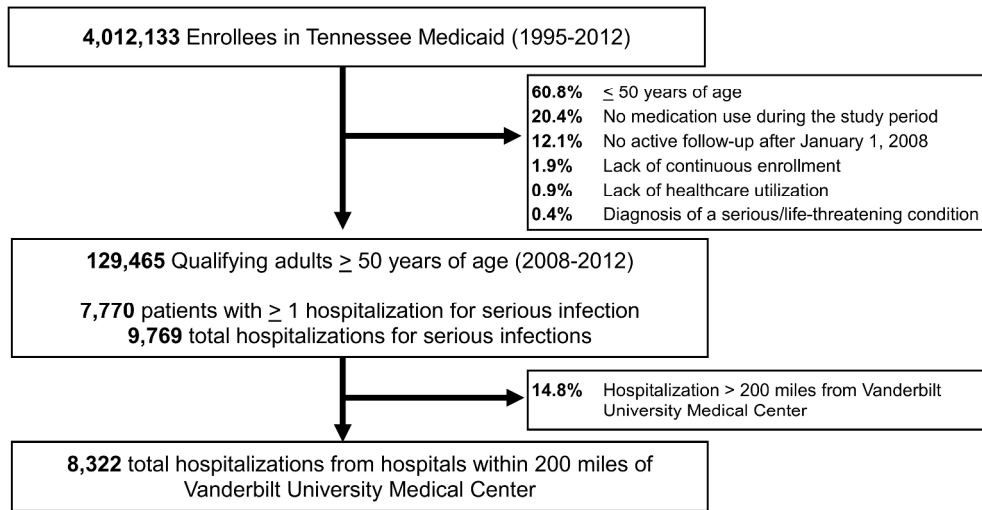


Figure 2. Identifying a retrospective cohort of patients ≥ 50 years of age without serious/life-threatening conditions, Tennessee Medicaid (2008-2012)

355x190mm (300 x 300 DPI)

Review only

Supplementary Appendix

Infection-Specific Definitions of Hospitalization for Serious Infection

We used a pre-specified adjudication process to determine whether each abstracted medical record corresponded to a true infection or not. Previous validation studies and expert clinical knowledge were used to define specific a priori definitions for each infection type.¹⁻³ Information abstracted from the medical record was compared to these *a priori* definitions for each infection type to make the final determination of whether a hospitalization represented a true infection or not.

Outline

I. Sepsis/Septicemia/Bacteremia/Septic Shock/Generalized Infection	Page 2
II. Pneumonia	Page 3
III. Cellulitis/Soft-tissue infections	Page 5
IV. Endocarditis.....	Page 6
V. Meningitis/Encephalitis	Page 7
VI. Pyelonephritis.....	Page 9
VII. Septic Arthritis/Osteomyelitis	Page 10
VIII. References.....	Page 11

I. Sepsis/Septicemia/Bacteremia/Septic Shock/Generalized Infection

Either of the following [1 or 2]:

1. Positive culture of a non-contaminant pathogen

i. Positive blood culture [any of the following (1-2)]

1. Any gram-negative organism, except:
 - a. No predominant organism
2. A gram positive organism, except:
 - a. Coagulase-negative *Staphylococcus*
 - b. *Bacillus spp.* (other than *Bacillus anthracis*)
 - c. *Corynebacterium spp.*
 - d. *Propionibacterium spp.*
 - e. *Micrococcus*
 - f. Diptheroids
 - g. Viridians Group Streptococci
 - h. Enterococci
 - i. *Clostridium perfringens*
 - j. *Aerococcus*
 - k. *Alcaligenes faecalis*
 - l. *Citrobacter*
 - m. *Neisseria subflava*
 - n. *Stomatococcus*
 - o. *Streptococcus bovis*
 - p. *Veillonella candidemia*
 - q. *Mycobacterium tuberculosis*
 - r. *S. salivarius*
 - s. "Gram Positive"
 - t. "No predominant organism"
 - u. *Streptococcus alpha*

2. At least two of the following, documented at admission +/- 2 days [i-iii]

i. Hypotension

1. Systolic BP \leq 90 mmHg
2. Reduction of systolic BP of 40mmHg from earliest measurement collected during the admission of interest

ii. Two of the following [1-4]:

1. Temperature \geq 38⁰C **or** \leq 36⁰C
2. Heart rate \geq 90 beats/minute
3. Respiratory rate \geq 20 breaths/min or PaCO₂ < 32 mmHg
4. WBC \geq 10,000 cells/mm³ **or** \leq 4,500 cells/mm³ **or** WBC with > 10 % immature (band) forms

iii. Initiation of antibiotic treatment specifically for

sepsis/septicemia/bacteremia/septic shock/generalized infection

II. Pneumonia

1. Pneumonia identified through examination (**all three of the following [a-c]**):
 - a. One of the following admission findings indicative of respiratory findings:
 1. New and/or increased cough
 2. Shortness of breath
 3. Pleuritic chest pain
 4. New purulent production
 5. Altered mental status (“agitation” and “lethargy” included)
 6. Crackles
 - a. Physical evidence of consolidation such as egophony, whispered pectoriloquy, etc.
 - b. One of the following examination findings indicative of systemic infection [1-4]:
 1. Temperature ($T \geq 100.4^{\circ}\text{F}$ (38°C) or $\leq 96^{\circ}\text{F}$) in first 48 hours of admission
 2. Systolic BP $\leq 90\text{mmHg}$
 3. Shock
 - a. Volume nonresponsive hypotension
 4. Blood peripheral WBC ($\geq 10.0 \times 10^9/\text{L}$ or $\leq 4.5 \times 10^9/\text{L}$)
 - c. Treatment with antibiotics/antivirals indicated for suspected infection

OR

At least two of the following [1-3]:

1. Two of the following from #1 (**[a and b], [a and c], or [b-c]**)
2. Any of the following findings listed on chest imaging from radiologic report **documented at admission +/- 2 days**
 - a. Pneumonia
 - b. Lung abscess
 - c. Opacity consistent with pneumonia/lung abscess
 - d. Infiltrate consistent with pneumonia/lung abscess
 - e. Consolidation consistent with pneumonia/lung abscess
 - f. Increased density consistent with pneumonia/lung abscess
 - g. Pleural effusion consistent with pneumonia/lung abscess
 - h. Interstitial edema consistent with pneumonia/lung abscess
3. Sterile Site Laboratory Findings
 - i. Any one of the following [i through v]
 - i. Sputum lab findings [any **one** of the following (1, 2)]:
 1. Sputum culture/PCR/serology/gram stain positive for an agent that is not considered a contaminant [see exclusion list below]:
 - a. *Aspergillus* species, *Enterococcus* species, viridians group streptococci, and yeast
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - ii. Blood lab findings [either of the following (1-3)]
 1. Blood culture/PCR/serology positive for an agent that is not considered a contaminant [see exclusion list below]:
 - a. Exclusions
 - i. Coagulase-negative *Staphylococcus*

- ii. *Bacillus spp.* (other than *Bacillus anthracis*)
 - iii. *Corynebacterium spp.*
 - iv. *Propionibacterium spp.*
 - v. *Micrococcus*
 - vi. Diphtheroids
 - vii. Viridians Group Streptococci
 - viii. Enterococci
 - ix. *Clostridium perfringens*
 - x. *Aerococcus*
 - xi. *Alcaligenes faecalis*
 - xii. *Citrobacter*
 - xiii. *Neisseria subflava*
 - xiv. *Stomatococcus*
 - xv. *Streptococcus bovis*
 - xvi. *Veillonella candidemia*
 - xvii. *Mycobacterium tuberculosis*
 - xviii. *S. salivarius*
2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
- iii. Pleural fluid lab findings [either of the following (1, 2)]
 1. Culture/PCR/serology positive for a bacterial pathogen
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - iv. Bronchoscopic specimen or deep endotracheal tube aspiration lab findings [either of the following (1, 2)]
 1. Culture/PCR/serology positive for a bacterial pathogen
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - v. Urine antigen detection testing [either of the following (1, 2)]
 1. *Legionella pneumophila*
 2. *Streptococcus pneumoniae*

III. Cellulitis/Soft-Tissue Infection

Both of the following:

1. Any mention of the following with recent onset (≤ 14 days) [*any of the following*]
 - a. Skin erythema
 - b. Surgical site infection
 - c. Superficial central line infection
 - d. Ostomy site infection
 - e. Skin infection with associated lymphangitis
2. Antibiotic treatment initiated for suspected infection

For peer review only

IV. Endocarditis

Any one of the following [1-3]:

1. Major Criteria [both of the following]:
 - a. Suggestive microbiology [at least one of the following]:
 - i. Positive blood culture of an *endocarditis organism* [any of the following]:
 1. *Streptococcus bovis*
 2. *Viridians streptococci*
 3. *Staphylococcus aureus*
 4. *Enterococcus spp.*
 5. HACEK organisms
 6. Coagulase negative staphylococci
 - b. Evidence of endocardial involvement [at least one of the following]:
 - i. New regurgitant murmur (a change in a preexisting murmur does not get scored)
 - ii. Echocardiogram suspicious for any of the following:
 1. Intracardiac mass with no alternative explanation
 2. Endocardial abscess
 3. New partial prosthesis dehiscence
 4. Vegetation on valve
2. Minor Criteria [at least 4 of the following]:
 - a. Predisposing valvular disease or IV drug use
 - b. Temperature $\geq 100.4^{\circ}\text{F}$ or 38°C
 - c. Vascular phenomena
 - i. Janeway lesions, conjunctival hemorrhages, arterial emboli, septic pulmonary infarcts, mycotic aneurysm, intracranial bleed
 - d. Immunologic phenomena
 - i. Osler nodes, Roth Spots, elevated Rheumatoid factor, hematuria in non-catheter urine, or other evidence of glomerulonephritis
 - e. Positive blood cultures
 - i. Excluding a single positive culture for coagulase negative staphylococci or a single positive culture for an organism that does not fall into the “reasonable endocarditis organism” (i.e. coagulase-positive and coagulase-negative *S. aureus*, *Enterococcus*, *viridians* group *Streptococci*, *S. bovis*, HACEK organisms)
 - f. Positive serology for *Brucella*, *Bartonella*, *Legionella*, or *Chlamydia*
 - g. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
3. At least one Major Criteria **AND** 3 minor criteria.

V. Meningitis/Encephalitis

Any one of the following [1 or 2]:

1. Both of the following [a-b]
 - a. Laboratory Findings [any one of the following (i-ix)]
 - i. CSF demonstrates any bacterium
 1. Excluding Diptheroids, Propionibacteria, Bacillus, Coagulase Negative *Staphylococcus*
 - ii. CSF demonstrates Diptheroids, Propionibacteria, Bacillus, Coagulase Negative *Staphylococcus* in the setting of past neurosurgical intervention **AND** physicians elected to treat with antibacterials
 - iii. Blood cultures positive for any of the following:
 1. *S. pneumoniae*
 2. *H. influenza*
 3. *Neisseria meningitidis*
 4. Group B Streptococcus
 - iv. Stool cultures positive for enterovirus
 - v. Throat or sputum cultures positive for *Neisseria meningitidis* in the setting of a rapid onset, overwhelming infection syndrome, including petechiae
 - vi. Serology positive for *Mycoplasma*, *Leptospira*, measles, mumps, lymphocytic choriomeningitis virus, arboviruses (e.g. St. Louis encephalitis virus), or HIV (if historically consistent with acute seroconversion).
 - vii. Brain biopsy demonstrates encephalitis
 - viii. Positive CSF culture or PCR detection for any of the following
 - ix. Acute or convalescent serology demonstrates positive antibody pattern for any of the following:
 1. Encephalitis arbovirus (La Crosse, St. Louis, Eastern Equine, Western Equine, Powassan, Japanese, West Nile)
 - b. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected meningitis/encephalitis
2. At least two of the following [a-d]
 - a. Clinical meningitis/encephalitis [at least two of the following]:
 - i. Petechial rash
 - ii. Nuchal rigidity (by history or exam)
 - iii. Altered sensorium
 - iv. Fever
 - v. Altered level of consciousness, including “agitation” or “lethargy”
 - vi. Behavioral change
 - vii. Diminished level of consciousness (not easily roused)
 - viii. History of any of the following: headaches, altered mental status, or recent exposure to patient with known bacterial meningitis
 - ix. Reduction in fever within 72 hours of starting anti-bacterial
 - b. Inflammatory CSF [at least one of the following i-ii]
 - i. Pleocytosis: ≥ 15 WBC/mm³ (after subtracting one WBC for every 1,000 RBC)
 - ii. Elevated protein (based on local lab-determined upper limits)
 - c. Suggestive Findings [at least one of the following (i-iv)]

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
- i. Septic syndrome
 - ii. Focal neurological deficits documented during examination (such as flaccid paralysis or speech alterations for West Nile Virus)
 - iii. Abnormal imaging
 1. Computed tomography or magnetic resonance imaging (MRI) demonstrating focal edema or inflammation or hemorrhage
 2. Indicated as “meningitis/encephalitis” or “compatible with meningitis/encephalitis” or “cannot rule out meningitis/encephalitis”
 - iv. Findings indicating an abnormal electroencephalography (such as focal periodic discharges)
- d. Antibiotic/antiviral/antifungal treatment initiated/recommended for presumed meningitis/encephalitis

VI. Pyelonephritis

At least two of the following [1-4]:

1. Suggestion of infection [at least one of the following]:
 - a. Temperature \geq 100.4°F (38°C)
 - b. Peripheral blood WBC \geq 10,000/mm³
 - c. Positive blood culture for any of the following:
 - i. Gram Negative Rods
 - ii. *Enterococcus spp.*
 - iii. *Staphylococcus saprophyticus*
 - d. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
2. Strong renal localization [at least one of the following]:
 - a. CT, MRI, or Ultrasound Suggestive of Renal Inflammation
3. Minor Criteria [at least two of the following]:
 - a. Flank pain
 - b. Costovertebral angle tenderness
 - c. Complaints of dysuria, frequency, or suprapubic pain
 - d. Any pyuria
 - e. Urine culture positive for a single organism
4. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected pyelonephritis

VII. Septic Arthritis/Osteomyelitis

Any one of the following (1-5):

1. Synovial fluid gram stain or tissue gram stain or special stain demonstrating any organism
2. Joint culture/PCR/serology positive for any organism
3. At least two of the following (a-d):
 - a. Positive blood culture/PCR/serology
 - b. Joint with acute (≤ 7 days) worsening of inflammatory features (**at least two of the following**):
 - i. Pain on history
 - ii. ROM
 - iii. Warmth
 - iv. Effusion
 - v. Swelling
 - vi. Limited range of motion
 - c. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
 - d. Any one of the following (i-iv)
 - i. Synovial fluid WBC $\geq 30,000/\text{mm}^3$
 - ii. Synovial fluid WBC $\geq 60,000/\text{mm}^3$ with $> 75\%$ PMNs
 - iii. Skin lesions, tenosynovitis, or urethral/cervical/rectal Gram stain or culture suggestive of *Neisseria gonorrhoeae*
 - iv. Any indication of the following in the synovial fluid: needle-like crystals, CPPD crystals, uric acid.
4. Positive bone biopsy [at least one of the following (a-c)]:
 - a. Positive culture for any organism
 - b. Positive gram stain
5. Imaging and indirect features [**at least two of the following (a-c)**]:
 - a. Consistent imaging [at least one of the following (i-iv)]:
 - i. Plain X-ray read by a radiologist as suggestive of osteomyelitis
 - ii. CT Scan read by a radiologist as suggestive of osteomyelitis
 - iii. MRI read by a radiologist as suggestive of osteomyelitis
 - iv. Bone scan or WBC scan read as suggestive of osteomyelitis
 - b. Suggestive indirect features[at least one of the following (i-viii)]:
 - i. Temperature $> 100.4^\circ\text{F}$ (38°C)
 - ii. Bony pain or tenderness or erythema over bone suspected to be infected
 - iii. Draining soft tissue sinus over bone suspected to be infected
 - iv. Positive “probe to bone” (or visible bone in deep ulcer at suspected site)
 - v. Blood culture positive for *S. aureus*
 - vi. ESR ≥ 75 mm/hour
 - vii. Intravenous drug use or indwelling catheter
 - viii. Inflammation on imaging associated with an orthopedic prosthesis
 - c. Positive culture for any organism from wound sample over the bone suspected of infection
 - d. Antibiotic/antiviral/antifungal treatment for suspected infection

VIII. References

1. Grijalva CG, Chung CP, Stein CM, et al. Computerized definitions showed high positive predictive values for identifying hospitalizations for congestive heart failure and selected infections in Medicaid enrollees with rheumatoid arthritis. *Pharmacoepidemiology and drug safety* 2008; 17(9): 890-5.
2. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital discharge databases coded serious bacterial infections accurately. *Journal of clinical epidemiology* 2007; 60(4): 397-409.
3. Patkar NM, Curtis JR, Teng GG, et al. Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients. *Journal of clinical epidemiology* 2009; 62(3): 321-7, 7.e1-7.

Section & Topic	No	Item	Reported on page #
Title or Abstract			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
<i>Participants</i>	6	Eligibility criteria	4,5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4,5
	9	Whether participants formed a consecutive, random or convenience series	4-6
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5, Table 1, Supplementary Appendix
	10b	Reference standard, in sufficient detail to allow replication	6, Supplementary Appendix
	11	Rationale for choosing the reference standard (if alternatives exist)	6, Supplementary Appendix
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6, Supplementary Appendix
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6, Supplementary Appendix
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6, Supplementary Appendix
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	7,8
	15	How indeterminate index test or reference standard results were handled	7,8
	16	How missing data on the index test and reference standard were handled	7,8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7,8
	18	Intended sample size and how it was determined	7,8
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	8
	20	Baseline demographic and clinical characteristics of participants	8
	21a	Distribution of severity of disease in those with the target condition	n/a
	21b	Distribution of alternative diagnoses in those without the target condition	n/a
	22	Time interval and any clinical interventions between index test and reference standard	n/a
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	9,10, Table 2
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	9,10 Table 2
	25	Any adverse events from performing the index test or the reference standard	n/a
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	11-13
	27	Implications for practice, including the intended use and clinical role of the index test	11-13

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60**OTHER
INFORMATION**

28	Registration number and name of registry	n/a
29	Where the full study protocol can be accessed	14
30	Sources of funding and other support; role of funders	13

For peer review only

BMJ Open

Validation of discharge diagnosis codes to identify serious infections among middle age and older adults

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020857.R2
Article Type:	Research
Date Submitted by the Author:	02-May-2018
Complete List of Authors:	Wiese, Andrew; Vanderbilt University Medical Center, Health Policy Griffin, Marie R; Vanderbilt University Medical Center, Health Policy Stein, Michael; Vanderbilt University, Pharmacology Schaffner, William; Vanderbilt University Medical Center, Health Policy Greevy, Robert; Vanderbilt School of Medicine, Biostatistics Mitchel, Jr., Edward; Vanderbilt University Medical Center, Health Policy Grijalva, Carlos; Vanderbilt University, Health Policy
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Infectious diseases, Epidemiology
Keywords:	coding algorithms, Medicaid, older adults, serious infections

SCHOLARONE™
Manuscripts

Peer Review Only

1
2
3 1 **Validation of discharge diagnosis codes to identify serious infections among middle age and older**
4
5 2 **adults**
6
7 3
8
9 4

10 **Running title:** Validation of diagnosis codes to identify infections
11
12 5

13
14 6 **Authors:** Andrew D. Wiese, PhD, MPH¹; Marie R. Griffin^{1,2}, MD, MPH; C. Michael Stein, MB, ChB³;
15
16 7 William Schaffner, MD¹; Robert Greevy, PhD⁴; Edward F. Mitchel Jr., MS¹; Carlos G. Grijalva, MD,
17
18 8 MPH^{1,2}
19

20 9 **Affiliations:** ¹Department of Health Policy, Vanderbilt University School of Medicine, Nashville,
21
22 10 Tennessee, USA; ²Mid-South Geriatric Research Education and Clinical Center, VA Tennessee Valley
23
24 11 Health Care System, Nashville, Tennessee, USA; ³Departments of Pharmacology and ⁴Biostatistics,
25
26 12 Vanderbilt University School of Medicine, Nashville, Tennessee, USA
27
28 13
29

30 14 **Corresponding Author:** Andrew D. Wiese, PhD, MPH; Department of Health Policy, Vanderbilt
31
32 15 University Medical Center, Suite 2600, Village at Vanderbilt, 1500 21st Avenue South, Nashville, TN
33
34 16 37212; phone: (615) 875-7997; email: andrew.d.wiese.1@vumc.org
35
36 17
37
38

39 18 **Key words:** coding algorithms; Medicaid; older adults; serious infections
40
41 19
42

43 20 **Word Count:** 3,777/4,000
44

45 21 **Tables and Figures:** (4/5)
46
47 22
48
49 23
50
51 24
52
53 25
54
55 26
56
57
58 1
59
60

1
2
3 27 **ABSTRACT (277/300)**
4

5 28 **Objectives:** Hospitalizations for serious infections are common among middle age and older adults and
6
7 29 frequently used as study outcomes. Yet few studies have evaluated the performance of diagnosis codes to
8
9 30 identify serious infections in this population. We sought to determine the positive predictive value (PPV)
10
11 31 of diagnosis codes for identifying hospitalizations due to serious infections among middle age and older
12
13 32 adults.

14
15 33 **Setting and participants:** We identified hospitalizations for possible infection among adults ≥ 50 years
16
17 34 enrolled in the Tennessee Medicaid healthcare program (2008-2012) using ICD-9 diagnosis codes for
18
19 35 pneumonia, meningitis/encephalitis, bacteremia/sepsis, cellulitis/soft-tissue infections, endocarditis,
20
21 36 pyelonephritis and septic arthritis/osteomyelitis.

22
23 37 **Design:** Medical records were systematically obtained from hospitals randomly selected from a stratified
24
25 38 sampling framework based on geographical region and hospital discharge volume.

26
27 39 **Measures:** Two trained clinical reviewers used a standardized extraction form to abstract information
28
29 40 from medical records. Pre-defined algorithms served as reference to adjudicate confirmed infection-
30
31 41 specific hospitalizations. We calculated the PPV of diagnosis codes using confirmed hospitalizations as
32
33 42 reference. Sensitivity analyses determined the robustness of the PPV to definitions that required
34
35 43 radiological or microbiological confirmation. We also determined interrater reliability between reviewers.

36
37 44 **Results:** The PPV of diagnosis codes for hospitalizations for infection ($n=716$) was 90% (95% CI: 88-
38
39 45 92). The PPV was highest for pneumonia [97% (95% CI: 94-98)] and cellulitis [91% (95% CI: 85-95)],
40
41 46 and lowest for meningitis/encephalitis [50% (95% CI: 24-76)]. The adjudication reliability was excellent
42
43 47 [93% agreement; first agreement-coefficient: 0.91]. The overall PPV was lower when requiring
44
45 48 microbiological confirmation [45%] and when requiring radiological confirmation for pneumonia [79%].

46
47 49 **Conclusions:** Discharge diagnosis codes have a high PPV for identifying hospitalizations for common,
48
49 50 serious infections among middle age and older adults. PPV estimates for rare infections were imprecise.

50
51 51

52
53 52

53 STRENGTHS AND LIMITATIONS OF THE STUDY

- 54 • This study examined the performance of diagnosis coding algorithms to identify hospitalizations
55 due to serious infections among middle age and older adults enrolled in a State Medicaid program
56 using a systematic and representative sample of records from hospitals of different sizes and in
57 distinct State regions.
- 58 • The reference criteria to identify true infections was based on the previous literature and clinical
59 expertise but may be imperfect. Nevertheless, identifying microbiologically-confirmed infections
60 is difficult due to the low sensitivity of culture-based diagnostic methods often used in clinical
61 practice.
- 62 • Diagnosis codes were based on the ICD-9-coding system only. These findings will continue to be
63 helpful for retrospective studies that encompass periods of ICD-9 use, yet additional studies
64 evaluating the performance of ICD-10-based codes would be beneficial.
- 65 • Our coding algorithms to identify serious infections had a high positive predictive value overall,
66 and will be useful in ongoing and future research using administrative data

78 INTRODUCTION

79 Infectious diseases remain a leading cause of morbidity and mortality in the U.S. and
80 elsewhere.(1) Middle age and older adults, in particular, are at high risk for serious infections and their
81 long-term consequences.(2, 3) Among older adults, community-acquired serious infections (including
82 pneumonia, sepsis, and meningitis) often require hospitalization and represent a substantial burden on the
83 U.S. healthcare system.(4-7) The incidence of community-acquired pneumonia is very high among adults
84 ≥ 50 years of age (248 cases per 100,000 adults) with an even higher burden among adults ≥ 80 years of
85 age (1,643 cases per 100,000 adults).(8) Sepsis, cellulitis and pyelonephritis are also very common
86 (sepsis: 100 cases per 100,000 and cellulitis/pyelonephritis: >150 hospitalizations per 100,000 adults)
87 with an increasing incidence of severe sepsis with increased age.(9-11) Meningitis and endocarditis are
88 relatively rare (around 2-3 cases per 100,000), although the case fatality rate is very high.(12, 13)
89 Therefore, it is important to monitor the incidence of these infections, identify important risk factors, and
90 determine the impact of preventative policies (e.g., vaccination) on these diseases among middle age and
91 older adults.(14-16)

92 Large-scale epidemiological studies using administrative data often use serious infections as
93 outcomes.(17-21) However, few studies have evaluated the performance of diagnosis codes to identify
94 serious infections among middle age and older adults. Most previous studies that have assessed the
95 performance of coded discharge diagnosis codes to identify serious infections have focused mainly on
96 common infections (e.g., pneumonia or sepsis), specific populations (e.g., patients with rheumatoid
97 arthritis), or on healthcare-associated or hospital-acquired infections.(22-31) Nevertheless, the
98 performance of coded discharge diagnoses for accurately identifying infections requiring hospitalization
99 among middle age and older adults is unclear. Therefore, we sought to determine the positive predictive
100 value (PPV) of specific discharge diagnoses for identifying infections that required hospitalization among
101 middle age and older adults.

102 103 METHODS

104 **Data sources**

105 TennCare is the managed Medicaid program in the State of Tennessee that provides healthcare
106 insurance to those who are Medicaid eligible (around 20% of the Tennessee population).(32) The adult
107 TennCare population consists of low-income pregnant women and individuals who are elderly or have a
108 disability (over 600,000 annually).(32) We used data from TennCare, supplemented with data from the
109 Tennessee Hospital Discharge Data System (a registry for all hospitalizations in Tennessee) and
110 pharmacy information from Medicare Part D for those that were dual eligible, to identify a retrospective
111 cohort of TennCare enrollees ≥ 50 years of age with pharmacy benefits (2008-2012). We restricted the
112 hospitalizations for serious infection to those occurring from 2008 through 2012 to only include more
113 recent hospitalizations for which medical records are more likely to be available. Cohort members had at
114 least 180 days of baseline continuous enrollment before cohort entry, and were also required to be free of
115 certain life-threatening conditions known to increase the risk of infection (solid organ transplantation,
116 end-stage renal disease, HIV/AIDS, malignancy and serious kidney, liver and respiratory disease) that
117 may limit longitudinal follow-up and impact the assessments of patients' exposures and their risk of
118 infections. Cohort members were also required to have evidence of at least one pharmacy prescription fill
119 and evidence of at least one healthcare encounter during baseline (to ensure use of benefits so that if a
120 healthcare encounter for an infection occurred, it would be detected). Follow-up started on the earliest
121 date the inclusion criteria were met and continued through the earliest of the following: study end date
122 (December 31, 2012), the day prior to diagnosis of a serious life-threatening condition that would have
123 precluded entry to the cohort, loss of enrollment, or date of death. From this retrospective cohort, we
124 identified possible hospitalizations for serious infections (see *Identification of hospitalizations for serious*
125 *infection*) for our validation study. To avoid including infections that may have originated due to a
126 previous hospital stay, we excluded hospitalizations for infections that occurred in the 30-day period after
127 discharge from a previous hospitalization. The study was approved by the institutional review boards of
128 Vanderbilt University and the Tennessee Department of Health, and by the Division of TennCare.

129 **Identification of hospitalizations for serious infection**

1
2
3 130 Clinical knowledge and a literature review were used to identify primary discharge diagnosis
4
5 131 codes that have been used previously to identify specific serious infections that require hospitalization
6
7 132 (*study infections*), including pneumonia (alone or with a primary diagnosis of bacteremia/sepsis),
8
9 133 bacteremia/sepsis, pyelonephritis, meningitis/encephalitis, osteomyelitis/septic arthritis, endocarditis and
10
11 134 cellulitis.(31, 33-35) Specific International Classifications of Diseases-Clinical Modification, 9th-revision
12
13 135 (ICD-9-CM) diagnosis codes used to identify possible hospitalizations for each infection type are
14
15 136 presented in Table 1. As the objective of our study was to determine the PPV of coding algorithms to
16
17 137 identify serious infections that required hospitalization, we focused only on primary diagnoses of
18
19 138 infection to reduce the possibility of detecting concurrent infections that may not have led to
20
21 139 hospitalization or nosocomial infections that developed during the course of the hospitalization.(35)

22 140 **Sampling Strategy**

23
24
25 141 We used stratified random sampling to select a representative subset of study infection
26
27 142 hospitalizations from among all possible cases identified in the retrospective cohort from among hospitals
28
29 143 within 200 miles of Vanderbilt University Medical Center (VUMC). Since larger hospitals would be
30
31 144 over-represented in a purely random sampling, and because there may also be regional variability in
32
33 145 coding practices and infection prevalence, we constructed a sampling framework where hospitals were
34
35 146 stratified based on their geographic region in Tennessee (West, Central, and East), and tertiles of reported
36
37 147 discharge volume (Low, Medium, and High) during the study period.(36-38) From this sampling
38
39 148 framework, we randomly selected three hospitals from each of these nine sampling strata, and retrieved
40
41 149 their medical records for review and validation (Figure 1). This strategy, relative to a purely random
42
43 150 sample, ensured better representation of infections identified in smaller hospitals and those in more rural
44
45 151 regions of the State of Tennessee. If a hospital refused to participate, it was replaced by another hospital
46
47 152 randomly selected from the same sampling stratum.

48
49
50
51 153 The overall goal was to review and validate 675 hospitalizations for serious infection from 27
52
53 154 hospitals (25 hospitalizations for each of the 3 hospitals comprising a stratum, yielding 75 hospitalizations
54
55 155 for each of the 9 strata) (Figure 1). We conservatively assumed that up to 80% of records requested would

1
2
3 156 be available for review, and so we requested 32 records per hospital to receive an average of 25 records
4
5 157 from each (Figure 1). To ensure that we reviewed sufficient rare infections, we preferentially selected any
6
7 158 identified possible hospitalizations for meningitis/encephalitis, osteomyelitis/septic arthritis and
8
9 159 endocarditis from each hospital in the sample. We randomly selected the remaining set of possible
10
11 160 hospitalizations for other serious infections based on the proportional distribution of common infections
12
13 161 at each hospital (pneumonia, bacteremia/sepsis, pyelonephritis and cellulitis) until 32 infections were
14
15 162 identified. For hospitals with fewer than 32 infections during the study period, all infections were
16
17 163 requested.

19 20 164 **Abstraction of Medical Records**

21
22 165 Relevant clinical information was abstracted from the medical record (transfer notes, emergency
23
24 166 room summary, admission summary, physical/history, pharmacy, laboratory, microbiology, and radiology
25
26 167 information, and discharge summary) of each hospitalization with a primary discharge diagnosis code
27
28 168 indicative of infection using a standardized and customized REDCap electronic data capture instrument
29
30 169 hosted at Vanderbilt University.(39) As we were interested in infections that led to hospitalizations, we
31
32 170 focused our reviews on clinical, microbiological and radiological information from the 2 days prior to the
33
34 171 admission date through 2 days after admission to limit the possibility of identifying infections that
35
36 172 developed during the hospitalization (i.e. nosocomial infections). In preparation for this study, the case
37
38 173 report form was pilot-tested among a separate, convenience sample of 354 possible infections identified
39
40 174 in the cohort from 3 hospitals in the same city as Vanderbilt University. This separate sample of
41
42 175 hospitalizations was used only for pilot-testing the case report form, and was not included in the current
43
44 176 study. One trained medical reviewer abstracted the relevant information for all selected records using the
45
46 177 case report form. During the abstraction process, the lack of a particular finding in the medical record was
47
48 178 treated as a lack of evidence for that finding, and so no information was considered missing after
49
50 179 abstraction.

51 52 53 180 **Adjudication of Medical Records**

1
2
3 181 All records received were reviewed, abstracted and adjudicated. We made the final determination
4
5 182 of whether a hospitalization represented a confirmed infection or not using *a priori* definitions of clinical,
6
7 183 radiological, and/or microbiological findings compatible with infection for each infection type. Previous
8
9 184 validation studies and expert clinical knowledge were used to define these specific *a priori* definitions for
10
11 185 each infection type (*Supplementary appendix*). (31, 35, 40)

13 186 **Statistical analysis**

15
16 187 We calculated the PPV of the ICD-9-CM discharge diagnosis codes for identifying
17
18 188 hospitalizations for serious infection using the results of the *a priori* definitions applied to the information
19
20 189 abstracted from the medical records as the reference (the proportion of cases identified with discharge
21
22 190 diagnosis codes that were determined to be true cases after adjudication of the medical record
23
24 191 information). We calculated 95% confidence intervals for the PPV using Wilson's formula. (41)
25
26 192 Secondary analyses assessed the PPV for hospitalizations for serious infection across hospitals of
27
28 193 different sizes and in different geographical regions of Tennessee.

29
30 194 We also assessed the reliability of the abstraction process. A second trained medical reviewer
31
32 195 abstracted relevant information from a subset of selected records, which included all meningitis and
33
34 196 endocarditis records, and a random selection of 10% of each of the remaining infection types. Each
35
36 197 reviewer conducted the process independently and blinded from one another. For the subset of records
37
38 198 abstracted by both reviewers, inter-reviewer agreement for the adjudication of a true or mis-identified
39
40 199 infection was assessed using the Gwet's first agreement coefficient (AC_1). (42-44) Since Cohen's kappa
41
42 200 statistic can be unreliable when the prevalence of the event and the level of observer agreement are high
43
44 201 in the study sample, we used Gwet's AC_1 as a reliability measure unlikely to be affected by these
45
46 202 concerns. (44-46)

47
48
49 203 In planned sensitivity analyses, we first assessed the impact of excluding hospitalizations that
50
51 204 occurred after the individual was transferred from another healthcare facility, as initial documentation and
52
53 205 details of the infection could be missing or incomplete in the receiving hospital. (40) We also assessed the
54
55 206 impact on the PPV for all infections when requiring microbiological identification of a pathogen

207 (excluding common contaminants) from a sterile site within 2 days before or after the hospitalization
208 admission date. A final sensitivity analysis among hospitalizations for possible pneumonia assessed the
209 PPV when radiological evidence of pneumonia was required [i.e. pneumonia, opacity, or infiltrate
210 mentioned in a chest X-ray or computed tomography scan report] (*Supplementary appendix*). All analyses
211 were performed in Stata-IC, version 15.1 (College Station TX).

212 Patient and Public Involvement

213 No patients were involved in the development of the research question, the outcome measures, or the
214 design or conduct of the study. As we conducted a retrospective study using administrative data, we have
215 no plans to disseminate the results of the research to study participants.

217 RESULTS

218 *Cohort characteristics*

219 Among a retrospective cohort of 129,465 adults ≥ 50 years of age enrolled in TennCare, 9,769
220 hospitalizations for serious infection were identified during the study period (2008-2012) among 7,770
221 unique patients (Figure 2). Cohort members were primarily female (57.8%) with a median age of 54 years
222 (mean: 57 years; range: 50-110). For efficiency considerations, our medical chart review activities then
223 focused on hospitalizations for serious infection (n=8,322) that occurred at hospitals within 200 miles of
224 VUMC. Pneumonia, cellulitis and bacteremia/sepsis were the most common infections identified using
225 discharge diagnosis codes (54.3%, 20.5% and 18.4%, respectively), followed by pyelonephritis (3.8%)
226 and septic arthritis/osteomyelitis (2.5%). Fewer than 1% of hospitalizations were due to
227 meningitis/encephalitis (n=30) and endocarditis (n=18).

228 *Collection, review and adjudication of selected medical records*

229 Of the 27 hospitals that were initially selected for the sample, 21 (78%) were able to participate.
230 We selected 7 additional hospitals to replace the 6 non-participants to achieve the desired sample size,
231 including an additional small hospital in the East region due to a large number of unavailable records
232 from a single participating hospital.

233 We received 716 (88.6%) of 808 requested records from 28 participating hospitals [Table 2].
234 Record availability from participating hospitals was lower in medium size hospitals (81.8%) compared to
235 small (93.5%) and large hospitals (91.7%), but did not differ by geographic region. Record availability by
236 infection type was greater than 86% for all infection types, with the exception of hospitalizations for the
237 rare endocarditis cases (57.1%; only 4 of 7 cases).

238 The sample of hospitalizations for serious infection included patients who were primarily female
239 (63.6%), with a median age of 60 years (mean: 64 years; range: 50-101) at the time of hospitalization.
240 There was evidence of transfer from a prior healthcare facility for 21.8% of the hospitalizations for
241 serious infection [highest percentage of transfers for bacteremia/sepsis (38.5%) and pneumonia (25.1%)].
242 The most common healthcare facility source was a nursing home/skilled nursing facility (84.6%), but also
243 included group home sources (7.7%), other sources (4.5%) [assisted living facility, mental health center]
244 and another acute care hospital (3.2%). There was evidence of an emergency department visit within 7
245 days prior to admission date for the serious infection hospitalization in 4.8% of the records.

246 *Performance of discharge diagnosis codes*

247 A total of 646 [PPV: 90.2% (95% CI: 87.8-92.2)] of the hospitalizations for serious infection
248 identified using ICD-9-CM primary discharge diagnosis codes were confirmed by applying the *a priori*
249 definitions to the abstracted data. The PPV was highest for pneumonia and cellulitis [96.5% (95% CI:
250 93.9-98.0) and 91.1% (95% CI: 84.7-94.9), respectively], and was $\geq 75\%$ for bacteremia/sepsis,
251 pyelonephritis, septic arthritis/osteomyelitis, and endocarditis. The PPV was lowest for
252 meningitis/encephalitis [50.0% (95% CI: 23.7-76.3)], although the precision was limited due to a low
253 number of available records for review (Table 2). Among the 10 potential cases of
254 meningitis/encephalitis, 7 cases were meningitis/meningoencephalitis and 3 were encephalitis. The
255 respective PPVs for meningitis/meningoencephalitis and encephalitis were 71.4% (95% CI: 35.9-91.8)
256 and 0%, respectively.

257 When performance was evaluated across stratification sampling parameters, no apparent
258 differences were observed in the PPV for records from hospitals in different geographical regions of

1
2
3 259 Tennessee. Although the PPV was high for all three discharge volume groups, the PPV was significantly
4
5 260 lower in large hospitals [84.6% (95% CI: 79.6-88.5)] compared to smaller hospitals [93.9% (95% CI:
6
7 261 90.0-96.3); PPV difference: -9.3% (95% CI: -14.7, -3.9)] and medium hospitals [92.7% (95% CI: 88.6-
8
9 262 95.4); PPV difference: -8.1% (95% CI: -13.7, -2.6)] (Table 2). This was likely driven by the different
10
11 263 distributions in the types of infections selected for review in the hospital groups. Large hospitals had a
12
13 264 higher proportion of non-pneumonia infections (70.4%) compared to medium and small hospitals (49.4%
14
15 265 and 36.1%, respectively). Importantly, the PPV for pneumonia was similar in each discharge volume
16
17 266 group (range: 96.0 to 96.6%), whereas the PPV was smaller for non-pneumonia infections in large
18
19 267 hospitals (79.8%) compared to medium (88.7%) and small (89.2%) hospitals.

20
21
22 268 In the 82 records independently abstracted by two reviewers to assess reliability, there was 92.7%
23
24 269 (95% CI: 86.9-98.4) agreement for identifying true hospitalizations for serious infection. The inter-rater
25
26 270 agreement was also high when assessing reliability, independent of the outcome prevalence, with an AC₁
27
28 271 of 0.91 (95% CI: 0.84-0.99). Of the 6 discordant cases, 3 were meningitis/encephalitis (1
29
30 272 meningitis/meningoencephalitis and 2 encephalitis), with one each of bacteremia/sepsis, pyelonephritis
31
32 273 and septic arthritis. The main reason for a discrepancy between reviewers was whether or not treatment
33
34 274 for the infection of interest occurred within 2 days of the admission date, which was one of the major
35
36 275 criteria for adjudication (see *Supplementary appendix*).

37 276 *Sensitivity analyses*

38
39
40
41 277 The PPV was virtually unchanged when excluding the 21.8% of hospitalizations that occurred as
42
43 278 transfers from another healthcare facility [90.1% (95% CI: 87.4-92.3)]. Microbiological evidence of the
44
45 279 specific infection type was found in 47.6% of records, leading to reduced PPVs when requiring
46
47 280 microbiological evidence [45.4% (95% CI: 41.8-49.1)]. Microbiological evidence of infection was
48
49 281 highest in hospitalizations for suspected pyelonephritis (94.4%), but was $\leq 60\%$ for every other infection
50
51 282 type [pneumonia (42.7%); cellulitis/soft tissue infections (58.5%); bacteremia/sepsis (26.1%)]. When
52
53 283 requiring radiological confirmation of pneumonia, the PPV for coded diagnoses was 78.8% (95% CI:
54
55 284 74.2-82.8). Approximately 95.6% of possible hospitalizations for pneumonia had at least one documented

1
2
3 285 chest x-ray or CT-scan. Among those patients with a chest x-ray or CT-scan report available (n=325),
4
5 286 83.4% had a finding compatible with pneumonia. The main findings among the 54 patients with possible
6
7 287 pneumonia and a radiological report available, but without radiological confirmation of pneumonia
8
9 288 included atelectasis (n=6), interstitial pneumonitis (n=3), chronic heart failure with pulmonary edema
10
11 289 (n=1), and no radiological findings of any kind (n=44).
12
13
14 290

15 291 **DISCUSSION**

16
17
18 292 Discharge diagnoses for identifying hospitalizations due to serious infections among middle age
19
20 293 and older adults had an overall PPV of 90.2%, with the highest values for the identification of common
21
22 294 serious infections. PPVs were consistently high across different hospital types and regions of Tennessee.
23
24 295 Furthermore, the PPV was similar after exclusion of hospitalizations for serious infection that were the
25
26 296 result of a transfer from another healthcare facility (e.g. acute care hospital, skilled nursing facility).
27
28 297 Microbiological confirmation was available for fewer than 50% of patients admitted with possible
29
30 298 hospitalizations for serious infections, and as expected, the requirement resulted in a low PPV for all
31
32 299 infections, with the exception of pyelonephritis. Importantly, the PPV for pneumonia hospitalizations
33
34 300 remained relatively high even when requiring radiological confirmation.
35
36

37 301 The PPV for hospitalizations for pneumonia in previous smaller validation studies has ranged
38
39 302 from 72 to 86% in different healthcare systems, but those studies were not focused on middle age and
40
41 303 older adults.(31, 47-49) In our study of hospitalizations among middle age and older adults, we found that
42
43 304 coded discharge diagnoses have a higher PPV for pneumonia compared to previous studies. The PPV for
44
45 305 bacteremia/sepsis was also on the higher range of previously reported PPVs for diagnosis codes to
46
47 306 identify bacteremia/sepsis from administrative data in other populations (reported range from 45% to
48
49 307 97.7%), and for septic arthritis/osteomyelitis compared to a previous study conducted among patients
50
51 308 with diabetes (63.9% versus 75.9% in our study).(23, 50, 51) Overall, the observed PPV for all infections
52
53 309 in our study was comparable to two previous comprehensive validation studies of bacterial infections, one
54
55 310 among patients with rheumatoid arthritis in a single hospital system and another among patients in one of
56
57

1
2
3 311 the Veteran's Affairs integrated service networks.(35, 40) Compared to these two previous studies of
4
5 312 ICD-9 codes, we abstracted and adjudicated a larger number of records while using a more systematic
6
7 313 sampling strategy to retrieve and review records for hospitalizations from multiple regions and hospital
8
9 314 types as opposed to a single hospital or healthcare system. However, some of the PPVs for individual
10
11 315 infections were less precise and less similar to these previous studies. This was especially true for rare
12
13 316 infections, as would be expected due to the low numbers of rare infections in our study and across
14
15
16 317 previous studies.(35, 40) The results of our study are also similar to previous validation studies that used
17
18 318 corresponding ICD-10 diagnosis codes to identify hospitalizations for serious infection.(52, 53)

19
20 319 One limitation to consider in our study was that it was not designed to estimate the sensitivity and
21
22 320 specificity of the coding algorithms. This would have required the identification, review and adjudication
23
24 321 of a sample of hospitalizations that did not fulfill our algorithm (i.e. presence of the ICD-9 primary
25
26 322 discharge diagnosis codes indicative of infection). However, when the prevalence of an outcome is low,
27
28 323 the PPV approximates the specificity.(54) Importantly, any non-differential outcome misclassification
29
30 324 between exposure groups resulting from the use of imperfect but highly-specific measurements would
31
32 325 attenuate the impact of the misclassification on the relative risk estimates.(55) In addition, we found that
33
34 326 the PPV of coded discharge diagnoses for serious infections remained high across hospitals of different
35
36 327 sizes and across different geographical areas of Tennessee, which may have different rates of
37
38 328 hospitalizations for serious infection.(56) Although our study applied a systematic sampling strategy to
39
40 329 assure the representation of different settings in our population, our population was restricted to middle
41
42 330 age and older adults enrolled in a State Medicaid program. Therefore, caution is warranted when
43
44 331 extrapolating the study findings to other populations.

45
46
47 332 Another limitation is the use of available clinical information to operationalize definitions for
48
49 333 adjudication of true hospitalizations for infections. It is possible that some procedures, laboratory findings
50
51 334 and diagnoses that informed the final diagnosis of infection were not fully recorded in the medical
52
53 335 records, and thus, were not available for our review and may have contributed to the observed PPV for
54
55 336 some infections. Although we used previous validation studies and clinical information to build pre-

1
2
3 337 specified definitions for the adjudication of true infections, our reference criteria may be imperfect,
4
5 338 considering the retrospective nature of our determinations and potential variability in clinical practice.
6
7 339 Nevertheless, we also assessed how the availability of selected findings (i.e. microbiological and
8
9 340 radiological information) in the medical record impacted the overall and infection-specific PPV. We
10
11 341 demonstrated that relying on highly specific clinical diagnostics, such as microbiological and radiological
12
13 342 information, to confirm true infections would result in lower PPVs for identification of infections in
14
15 343 administrative data. Requiring microbiological confirmation to confirm true infections is challenging
16
17 344 because of the known low sensitivity of culture-based diagnostic methods (most commonly used in
18
19 345 clinical practice), which may lead to misclassification.^(57, 58) In addition, requiring radiological evidence
20
21 346 compatible with pneumonia within 2 days of hospital admission did lower the observed PPV for
22
23 347 pneumonia hospitalizations. Nevertheless, the observed PPV remained close to 80%, which should reduce
24
25 348 concerns about using diagnosis codes to identify hospitalizations due to pneumonia. Finally, the coding
26
27 349 algorithms were based on the ICD-9-coding system only. Although these findings will be helpful for
28
29 350 retrospective studies that encompass periods of ICD-9 use, additional studies evaluating the performance
30
31 351 of ICD-10-based codes would be useful to complement our findings.

32
33
34
35 352 Our study demonstrated that discharge diagnosis codes can be used to accurately identify
36
37 353 hospitalizations for serious infections among middle age and older adults. The highest PPVs were
38
39 354 observed for the most common infections, and the PPV for pneumonia remained high when requiring
40
41 355 radiological confirmation. Importantly, consistently high PPVs were observed across different hospital
42
43 356 sizes and regions. However, the estimated PPV was lower and less precise for very rare infections (e.g.
44
45 357 encephalitis). This should be an important consideration for studies specifically focused on those less
46
47 358 frequent outcomes, especially when strict microbiological confirmation is required. Taken together, these
48
49 359 findings support the use of discharge diagnosis codes for infections to identify outcomes in ongoing and
50
51 360 future epidemiological studies among middle age and older adults.

361

362 **Acknowledgement**

14

1
2
3 363 We are indebted to the Tennessee Bureau of TennCare of the Tennessee Department of Finance and
4
5 364 Administration, which provided data for the study. We are also indebted to the Tennessee Department of
6
7 365 Health for providing data for the study. Statements in the report should not be construed as endorsement
8
9 366 by the U.S. Department of Health and Human Services, the Department of Veterans Affairs, or the
10
11 367 Tennessee Department of Health.
12
13
14 368

15 16 369 **Funding**

17
18 370 This study was funded by the NIH (R03-AG-042981 and R01-AG-043471-01A1) and the TL1 award
19
20 371 TL1TR000447.
21

22 372 **Contributors**

23
24 373 ADW planned the medical record collection and statistical analysis, analyzed and interpreted the data, and
25
26 374 drafted and revised the paper. MRG, WS, CMS, and RAG planned the statistical analysis, interpreted the
27
28 375 data and revised the paper. EFM prepared the data, and revised the paper. CGG designed the project,
29
30 376 acquired the data from TennCare, planned the medical record collection and statistical analysis,
31
32 377 interpreted the data, and revised the paper.
33

34 35 378 **Declaration of interests**

36
37 379 CGG has received consulting fees from Pfizer and Merck, and received research support from Sanofi-
38
39 380 Pasteur, Campbell Alliance, the Centers for Disease Control and Prevention, National Institutes of Health,
40
41 381 The Food and Drug Administration, and the Agency for Health Care Research and Quality. WS has
42
43 382 received personal fees from Pfizer, Merck, Novavax, Dynavax, Sanofi-Pasteur, GSK, Seqirus, and
44
45 383 received research support from the Centers for Disease Control and Prevention. ADW, EFM, CMS, MRG
46
47 384 and RAG have no conflicts of interest to disclose.
48

49 385 **Data sharing**

50
51 386 No additional unpublished data are available from the study. The study protocol and statistical code are
52
53 387 available from the corresponding author, Andrew Wiese (andrew.d.wiese.1@vumc.org).
54
55 388

389 **REFERENCES**

- 390 1. Khabbaz RF, Moseley RR, Steiner RJ, Levitt AM, Bell BP. Challenges of infectious diseases in
391 the USA. *Lancet*. 2014;384(9937):53-63.
- 392 2. Grohskopf LA, Sokolow LZ, Broder KR, Olsen SJ, Karron RA, Jernigan DB, et al. Prevention
393 and Control of Seasonal Influenza with Vaccines. *MMWR Recomm Rep*. 2016;65(5):1-54.
- 394 3. Kim DK, Bridges CB, Harriman KH. Advisory committee on immunization practices
395 recommended immunization schedule for adults aged 19 years or older--United States, 2015.
396 *MMWR Morb Mortal Wkly Rep*. 2015;64(4):91-2.
- 397 4. Jeon MH, Chung JW, Choi SH, Kim TH, Lee EJ, Choo EJ. Pneumonia risk factors and clinical
398 features of hospitalized patients older than 15 years with pandemic influenza A (H1N1) in South
399 Korea: a multicenter study. *Diagn Microbiol Infect Dis*. 2011;70(2):230-5.
- 400 5. Pilishvili T, Bennett NM. Pneumococcal Disease Prevention Among Adults: Strategies for the
401 Use of Pneumococcal Vaccines. *Am J Prev Med*. 2015;49(6 Suppl 4):S383-90.
- 402 6. Sjoding MW, Prescott HC, Wunsch H, Iwashyna TJ, Cooke CR. Longitudinal Changes in ICU
403 Admissions Among Elderly Patients in the United States. *Crit Care Med*. 2016;44(7):1353-60.
- 404 7. Crotty MP, Meyers S, Hampton N, Bledsoe S, Ritchie DJ, Buller RS, et al. Epidemiology, Co-
405 Infections, and Outcomes of Viral Pneumonia in Adults: An Observational Cohort Study.
406 *Medicine (Baltimore)*. 2015;94(50):e2332.
- 407 8. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM, et al. Community-Acquired
408 Pneumonia Requiring Hospitalization among U.S. Adults. *N Engl J Med*. 2015;373(5):415-27.
- 409 9. Filbin MR, Arias SA, Camargo CA, Jr., Barche A, Pallin DJ. Sepsis visits and antibiotic
410 utilization in U.S. emergency departments*. *Crit Care Med*. 2014;42(3):528-35.
- 411 10. Christensen KL, Holman RC, Steiner CA, Sejvar JJ, Stoll BJ, Schonberger LB. Infectious disease
412 hospitalizations in the United States. *Clin Infect Dis*. 2009;49(7):1025-35.

- 1
2
3 413 11. Ginde AA, Moss M, Shapiro NI, Schwartz RS. Impact of older age and nursing home residence
4
5 414 on clinical outcomes of US emergency department visits for severe sepsis. *J Crit Care.*
6
7 415 2013;28(5):606-11.
8
9 416 12. Lung B, Vahanian A. Epidemiology of acquired valvular heart disease. *Can J Cardiol.*
10
11 417 2014;30(9):962-70.
12
13 418 13. Thigpen MC, Whitney CG, Messonnier NE, Zell ER, Lynfield R, Hadler JL, et al. Bacterial
14
15 419 meningitis in the United States, 1998-2007. *N Engl J Med.* 2011;364(21):2016-25.
16
17 420 14. Simonetti AF, Garcia-Vidal C, Viasus D, Garcia-Somoza D, Dorca J, Gudiol F, et al. Declining
18
19 421 mortality among hospitalized patients with community-acquired pneumonia. *Clin Microbiol*
20
21 422 *Infect.* 2016;22(6):567.e1-7.
22
23 423 15. Rivero-Calle I, Pardo-Seco J, Aldaz P, Vargas DA, Mascaros E, Redondo E, et al. Incidence and
24
25 424 risk factor prevalence of community-acquired pneumonia in adults in primary care in Spain
26
27 425 (NEUMO-ES-RISK project). *BMC Infect Dis.* 2016;16(1):645.
28
29 426 16. Park JY, Park S, Lee SH, Lee MG, Park YB, Oh KC, et al. Microorganisms Causing Community-
30
31 427 Acquired Acute Bronchitis: The Role of Bacterial Infection. *PLoS One.* 2016;11(10):e0165553.
32
33 428 17. Jackson ML, Walker R, Lee S, Larson E, Dublin S. Predicting 2-Year Risk of Developing
34
35 429 Pneumonia in Older Adults without Dementia. *J Am Geriatr Soc.* 2016;64(7):1439-47.
36
37 430 18. Clement RC, Haddix KP, Creighton RA, Spang JT, Tennant JN, Kamath GV. Risk Factors for
38
39 431 Infection After Knee Arthroscopy: Analysis of 595,083 Cases From 3 United States Databases.
40
41 432 *Arthroscopy.* 2016;32(12):2556-61.
42
43 433 19. Weinert BA, Edmonson MB. Hospitalizations at Nonfederal Facilities for Lower Respiratory
44
45 434 Tract Infection in American Indian and Alaska Native Children Younger than 5 Years of Age,
46
47 435 1997-2012. *J Pediatr.* 2016;175:33-9.e4.
48
49 436 20. Akiyama T, Chikuda H, Yasunaga H, Horiguchi H, Fushimi K, Saita K. Incidence and risk
50
51 437 factors for mortality of vertebral osteomyelitis: a retrospective analysis using the Japanese
52
53 438 diagnosis procedure combination database. *BMJ Open.* 2013;3(3).

- 1
2
3 439 21. Graversen ME, Dalgaard LS, Jensen-Fangel S, Jespersen B, Ostergaard L, Sogaard OS. Risk and
4
5 440 outcome of pyelonephritis among renal transplant recipients. *BMC Infect Dis.* 2016;16:264.
6
7 441 22. Abou Zahr Z, Spiegelman A, Cantu M, Ng B. Perioperative use of anti-rheumatic agents does not
8
9 442 increase early postoperative infection risks: a Veteran Affairs' administrative database study.
10
11 443 *Rheumatol Int.* 2015;35(2):265-72.
12
13 444 23. Carnahan RM, Herman RA, Moores KG. A systematic review of validated methods for
14
15 445 identifying transfusion-related sepsis using administrative and claims data. *Pharmacoepidemiol*
16
17 446 *Drug Saf.* 2012;21 Suppl 1:222-9.
18
19 447 24. Condell O, Gubbels S, Nielsen J, Espenhain L, Frimodt-Moller N, Engberg J, et al. Automated
20
21 448 surveillance system for hospital-acquired urinary tract infections in Denmark. *J Hosp Infect.*
22
23 449 2016;93(3):290-6.
24
25 450 25. Gedeborg R, Furebring M, Michaelsson K. Diagnosis-dependent misclassification of infections
26
27 451 using administrative data variably affected incidence and mortality estimates in ICU patients. *J*
28
29 452 *Clin Epidemiol.* 2007;60(2):155-62.
30
31 453 26. Goto M, Ohl ME, Schweizer ML, Perencevich EN. Accuracy of administrative code data for the
32
33 454 surveillance of healthcare-associated infections: a systematic review and meta-analysis. *Clin*
34
35 455 *Infect Dis.* 2014;58(5):688-96.
36
37 456 27. Shaklee J, Zerr DM, Elward A, Newland J, Leckerman K, Asti L, et al. Improving surveillance
38
39 457 for pediatric *Clostridium difficile* infection: derivation and validation of an accurate case-finding
40
41 458 tool. *Pediatr Infect Dis J.* 2011;30(3):e38-40.
42
43 459 28. Wright SB, Huskins WC, Dokholyan RS, Goldmann DA, Platt R. Administrative databases
44
45 460 provide inaccurate data for surveillance of long-term central venous catheter-associated
46
47 461 infections. *Infect Control Hosp Epidemiol.* 2003;24(12):946-9.
48
49 462 29. Greenberg JA, Hohmann SF, Hall JB, Kress JP, David MZ. Validation of a Method to Identify
50
51 463 Immunocompromised Patients with Severe Sepsis in Administrative Databases. *Ann Am Thorac*
52
53 464 *Soc.* 2016;13(2):253-8.

- 1
2
3 465 30. Olsen MA, Young-Xu Y, Stwalley D, Kelly CP, Gerding DN, Saeed MJ, et al. The burden of
4 466 clostridium difficile infection: estimates of the incidence of CDI from U.S. Administrative
5 467 databases. *BMC Infect Dis.* 2016;16:177.
6
7
8
9 468 31. Grijalva CG, Chung CP, Stein CM, Gideon PS, Dyer SM, Mitchel EF, Jr., et al. Computerized
10 469 definitions showed high positive predictive values for identifying hospitalizations for congestive
11 470 heart failure and selected infections in Medicaid enrollees with rheumatoid arthritis.
12 471 *Pharmacoepidemiol Drug Saf.* 2008;17(9):890-5.
13
14
15
16 472 32. TennCare Fiscal Year 2015-2016 Annual Report. In: Administration DoHCF, ed; 2016.
17
18 473 33. Wiese AD, Griffin MR, Stein CM, Mitchel EF, Jr., Grijalva CG. Opioid Analgesics and the Risk
19 474 of Serious Infections Among Patients With Rheumatoid Arthritis: A Self-Controlled Case Series
20 475 Study. *Arthritis Rheumatol.* 2016;68(2):323-31.
21
22
23
24 476 34. Grijalva CG, Kaltenbach L, Arbogast PG, Mitchel EF, Jr., Griffin MR. Initiation of rheumatoid
25 477 arthritis treatments and the risk of serious infections. *Rheumatology (Oxford).* 2010;49(1):82-90.
26
27
28
29 478 35. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital
30 479 discharge databases coded serious bacterial infections accurately. *J Clin Epidemiol.*
31 480 2007;60(4):397-409.
32
33
34
35 481 36. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values
36 482 with disease prevalence. *Stat Med.* 1997;16(9):981-91.
37
38
39 483 37. Kim HM, Smith EG, Stano CM, Ganoczy D, Zivin K, Walters H, et al. Validation of key
40 484 behaviourally based mental health diagnoses in administrative data: suicide attempt, alcohol
41 485 abuse, illicit drug abuse and tobacco use. *BMC Health Serv Res.* 2012;12:18.
42
43
44
45 486 38. El-Ghitany EM, Farghaly AG, Farag S, Abd El-Wahab EW. Validation of EGCRISC for Chronic
46 487 Hepatitis C Infection Screening and Risk Assessment in the Egyptian Population. *PLoS One.*
47 488 2016;11(12):e0168649.
48
49
50
51
52
53
54
55
56
57

- 1
2
3 489 39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data
4 capture (REDCap)--a metadata-driven methodology and workflow process for providing
5 490 translational research informatics support. *J Biomed Inform.* 2009;42(2):377-81.
6 491
7
8 492 40. Patkar NM, Curtis JR, Teng GG, Allison JJ, Saag M, Martin C, et al. Administrative codes
9 combined with medical records based criteria accurately identified bacterial infections among
10 493 rheumatoid arthritis patients. *J Clin Epidemiol.* 2009;62(3):321-7, 7.e1-7.
11 494
12
13 495 41. Niesner K, Murff HJ, Griffin MR, Wasserman B, Greevy R, Grijalva CG, et al. Validation of VA
14 administrative data algorithms for identifying cardiovascular disease hospitalization.
15 496
16 497 *Epidemiology.* 2013;24(2):334-5.
17
18 498 42. Govatsmark RE, Sneeggen S, Karlsaune H, Slordahl SA, Bonna KH. Interrater reliability of a
19 national acute myocardial infarction register. *Clin Epidemiol.* 2016;8:305-12.
20 499
21
22 500 43. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and
23 Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with
24 501 personality disorder samples. *BMC Med Res Methodol.* 2013;13:61.
25 502
26
27 503 44. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br*
28 *J Math Stat Psychol.* 2008;61(Pt 1):29-48.
29 504
30
31 505 45. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: resolution of the two
32 paradoxes. *J Clin Epidemiol.* 1996;49(4):431-4.
33 506
34
35 507 46. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J*
36 *Clin Epidemiol.* 1990;43(6):543-9.
37 508
38
39 509 47. Drahos J, Vanwormer JJ, Greenlee RT, Landgren O, Koshiol J. Accuracy of ICD-9-CM codes in
40 identifying infections of pneumonia and herpes simplex virus in administrative data. *Ann*
41 510 *Epidemiol.* 2013;23(5):291-3.
42 511
43
44 512 48. Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients
45 with pneumonia. *Am J Med Qual.* 2005;20(6):319-28.
46 513
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 514 49. Kern DM, Davis J, Williams SA, Tunceli O, Wu B, Hollis S, et al. Validation of an
4
5 515 administrative claims-based diagnostic code for pneumonia in a US-based commercially insured
6
7 516 COPD population. *Int J Chron Obstruct Pulmon Dis*. 2015;10:1417-25.
8
9 517 50. Iwashyna TJ, Odden A, Rohde J, Bonham C, Kuhn L, Malani P, et al. Identifying patients with
10
11 518 severe sepsis using administrative claims: patient-level validation of the angus implementation of
12
13 519 the international consensus conference definition of severe sepsis. *Med Care*. 2014;52(6):e39-43.
14
15 520 51. Newton KM, Wagner EH, Ramsey SD, McCulloch D, Evans R, Sandhu N, et al. The use of
16
17 521 automated data to identify complications and comorbidities of diabetes: a validation study. *J Clin*
18
19 522 *Epidemiol*. 1999;52(3):199-207.
20
21 523 52. Sahli L, Lapeyre-Mestre M, Derumeaux H, Moulis G. Positive predictive values of selected
22
23 524 hospital discharge diagnoses to identify infections responsible for hospitalization in the French
24
25 525 national hospital database. *Pharmacoepidemiol Drug Saf*. 2016;25(7):785-9.
26
27 526 53. Holland-Bill L, Xu H, Sorensen HT, Acquavella J, Svaerke C, Gammelager H, et al. Positive
28
29 527 predictive value of primary inpatient discharge diagnoses of infection among cancer patients in
30
31 528 the Danish National Registry of Patients. *Ann Epidemiol*. 2014;24(8):593-7, 7.e1-18.
32
33 529 54. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic
34
35 530 research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-37.
36
37 531 55. Rodgers A, MacMahon S. Systematic underestimation of treatment effects as a result of
38
39 532 diagnostic test inaccuracy: implications for the interpretation and design of thromboprophylaxis
40
41 533 trials. *Thromb Haemost*. 1995;73(2):167-71.
42
43 534 56. van Walraven C, English S, Austin PC. Administrative database code accuracy did not vary
44
45 535 notably with changes in disease prevalence. *J Clin Epidemiol*. 2016;79:86-9.
46
47 536 57. Mancini N, Carletti S, Ghidoli N, Cichero P, Burioni R, Clementi M. The era of molecular and
48
49 537 other non-culture-based methods in diagnosis of sepsis. *Clin Microbiol Rev*. 2010;23(1):235-51.
50
51 538 58. Niederman MS. The argument against using quantitative cultures in clinical trials and for the
52
53 539 management of ventilator-associated pneumonia. *Clin Infect Dis*. 2010;51 Suppl 1:S93-9.
54
55
56
57
58 21
59
60

540 **Table 1. Discharge diagnosis code definitions (ICD-9-CM) for hospitalizations for serious infection**

Serious Infection	Primary (first listed) discharge diagnosis code
Pneumonia-primary definition	003.22, 480.* [†] , 481, 482.*, 483.*, 484.*, 485.*, 486.*, 487.0
Pneumonia-secondary definition (primary diagnosis code with pneumonia diagnosis (above) in any other diagnosis field)	510.*, 038.*, 790.7, 995.91, 995.92
Meningitis/ Encephalitis	003.21, 036.0, 047*, 049.*, 053.0, 054.72, 072.1, 091.81, 094.2, 098.82, 100.81, 320.*, 036.1, 054.3, 056.01, 058.21, 058.29, 062.*, 063.*, 064.*, 066.41, 072.2, 094.81, 130.0, 323.*
Bacteremia/ Sepsis [†]	038.*, 790.7, 995.91, 995.92
Cellulitis/ Soft-tissue infections	035, 040.0, 569.61, 681.*, 682.*, 728.86, 785.4
Endocarditis	036.42, 074.22, 093.2*, 098.84, 421.*, 422.92
Pyelonephritis	590.*
Septic Arthritis/ Osteomyelitis	003.23, 056.71, 098.5*, 711.0, 711.00-711.07, 711.09, 711.9*, 003.24, 376.03, 526.4, 730.0*, 730.1*, 730.2*

541 ‡ A * indicates all numeric values [0-9]

542 † Without a diagnosis of pneumonia in any other diagnosis field

543

Table 2. Positive predictive value (PPV) of coded discharge diagnosis definitions for hospitalizations for serious infections among adults ≥ 50 years of age enrolled in Tennessee Medicaid, 2008-2012

Type	Expected Number of Records	Records Received	PPV (95 % CI)
Overall	675	716	90.2 (87.8, 92.2)
Region Specific			
West	225	195	91.3 (86.5, 94.5)
Central	225	225	88.9 (84.1, 92.4)
East	225	296	90.5 (86.7, 93.4)
Bed volume size specific			
Low	225	230	93.9 (90.0, 96.3)
Medium	225	233	92.7 (88.6, 95.4)
High	225	253	84.6 (79.6, 88.5)
Serious Infection			
Pneumonia	305	340	96.5 (93.9, 98.0)
Cellulitis/Soft-tissue infections	125	123	91.1 (84.7, 94.9)
Pyelonephritis	80	89	87.6 (79.2, 93.0)
Bacteremia/Sepsis	100	92	82.6 (73.6, 89.0)
Septic Arthritis/Osteomyelitis	50	58	75.9 (63.5, 85.0)
Endocarditis	5	4	75.0 (30.1, 95.4)
Meningitis/Encephalitis	10	10	50.0 (23.7, 76.3)

Figure legends

Figure 1. Sampling strategy for identifying potential hospitalizations for serious infection

Figure 2. Identifying a retrospective cohort of patients ≥ 50 years of age without serious/life-threatening conditions, Tennessee Medicaid (2008-2012)

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

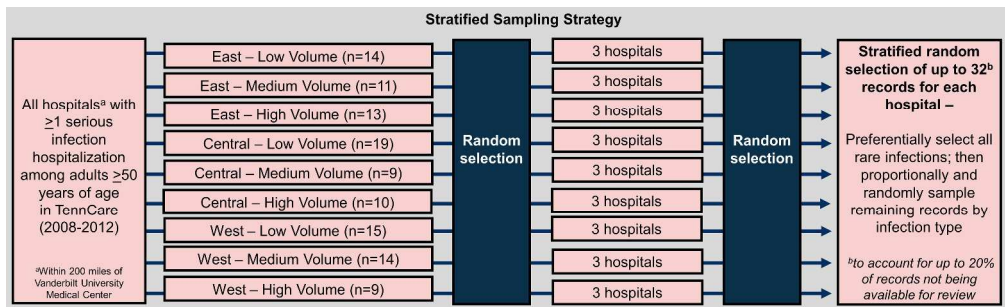


Figure 1. Sampling strategy for identifying potential hospitalizations for serious infection

457x139mm (300 x 300 DPI)

peer review only

BMJ Open: first published as 10.1136/bmjopen-2017-020857 on 19 June 2018. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

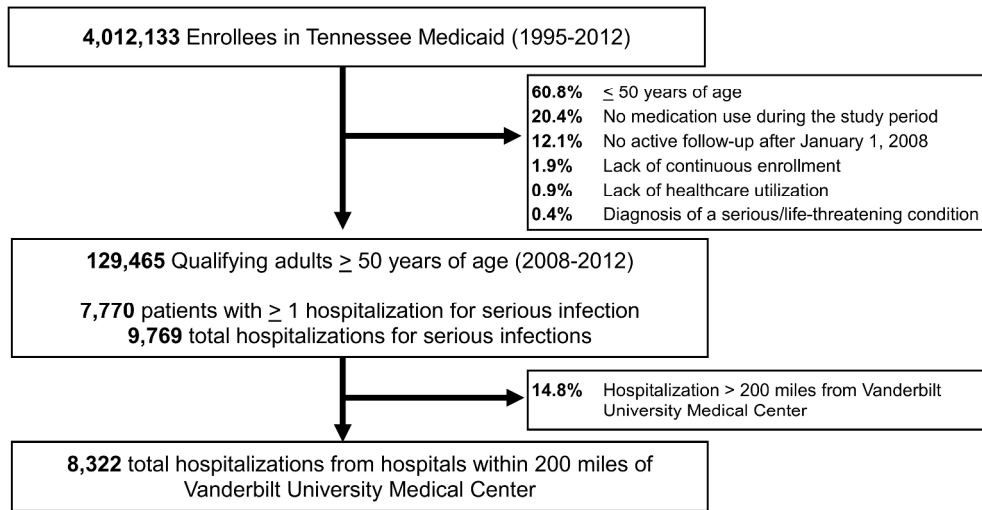


Figure 2. Identifying a retrospective cohort of patients ≥ 50 years of age without serious/life-threatening conditions, Tennessee Medicaid (2008-2012)

355x190mm (300 x 300 DPI)

Supplementary Appendix

Infection-Specific Definitions of Hospitalization for Serious Infection

We used a pre-specified adjudication process to determine whether each abstracted medical record corresponded to a true infection or not. Previous validation studies and expert clinical knowledge were used to define specific a priori definitions for each infection type.¹⁻³ Information abstracted from the medical record was compared to these *a priori* definitions for each infection type to make the final determination of whether a hospitalization represented a true infection or not.

Outline

I. Sepsis/Septicemia/Bacteremia/Septic Shock/Generalized Infection	Page 2
II. Pneumonia	Page 3
III. Cellulitis/Soft-tissue infections	Page 5
IV. Endocarditis.....	Page 6
V. Meningitis/Encephalitis	Page 7
VI. Pyelonephritis.....	Page 9
VII. Septic Arthritis/Osteomyelitis	Page 10
VIII. References.....	Page 11

I. Sepsis/Septicemia/Bacteremia/Septic Shock/Generalized Infection

Either of the following [1 or 2]:

1. Positive culture of a non-contaminant pathogen

i. Positive blood culture [any of the following (1-2)]

1. Any gram-negative organism, except:
 - a. No predominant organism
2. A gram positive organism, except:
 - a. Coagulase-negative *Staphylococcus*
 - b. *Bacillus spp.* (other than *Bacillus anthracis*)
 - c. *Corynebacterium spp.*
 - d. *Propionibacterium spp.*
 - e. *Micrococcus*
 - f. Diptheroids
 - g. Viridians Group Streptococci
 - h. Enterococci
 - i. *Clostridium perfringens*
 - j. *Aerococcus*
 - k. *Alcaligenes faecalis*
 - l. *Citrobacter*
 - m. *Neisseria subflava*
 - n. *Stomatococcus*
 - o. *Streptococcus bovis*
 - p. *Veillonella candidemia*
 - q. *Mycobacterium tuberculosis*
 - r. *S. salivarius*
 - s. "Gram Positive"
 - t. "No predominant organism"
 - u. *Streptococcus alpha*

2. At least two of the following, documented at admission +/- 2 days [i-iii]

i. Hypotension

1. Systolic BP \leq 90 mmHg
2. Reduction of systolic BP of 40mmHg from earliest measurement collected during the admission of interest

ii. Two of the following [1-4]:

1. Temperature \geq 38⁰C **or** \leq 36⁰C
2. Heart rate \geq 90 beats/minute
3. Respiratory rate \geq 20 breaths/min or PaCO₂ < 32 mmHg
4. WBC \geq 10,000 cells/mm³ **or** \leq 4,500 cells/mm³ **or** WBC with > 10 % immature (band) forms

iii. Initiation of antibiotic treatment specifically for

sepsis/septicemia/bacteremia/septic shock/generalized infection

II. Pneumonia

1. Pneumonia identified through examination (**all three of the following [a-c]**):
 - a. One of the following admission findings indicative of respiratory findings:
 1. New and/or increased cough
 2. Shortness of breath
 3. Pleuritic chest pain
 4. New purulent production
 5. Altered mental status (“agitation” and “lethargy” included)
 6. Crackles
 - a. Physical evidence of consolidation such as egophony, whispered pectoriloquy, etc.
 - b. One of the following examination findings indicative of systemic infection [1-4]:
 1. Temperature ($T \geq 100.4^{\circ}\text{F}$ (38°C) or $\leq 96^{\circ}\text{F}$) in first 48 hours of admission
 2. Systolic BP $\leq 90\text{mmHg}$
 3. Shock
 - a. Volume nonresponsive hypotension
 4. Blood peripheral WBC ($\geq 10.0 \times 10^9/\text{L}$ or $\leq 4.5 \times 10^9/\text{L}$)
 - c. Treatment with antibiotics/antivirals indicated for suspected infection

OR

At least two of the following [1-3]:

1. Two of the following from #1 (**[a and b], [a and c], or [b-c]**)
2. Any of the following findings listed on chest imaging from radiologic report **documented at admission +/- 2 days**
 - a. Pneumonia
 - b. Lung abscess
 - c. Opacity consistent with pneumonia/lung abscess
 - d. Infiltrate consistent with pneumonia/lung abscess
 - e. Consolidation consistent with pneumonia/lung abscess
 - f. Increased density consistent with pneumonia/lung abscess
 - g. Pleural effusion consistent with pneumonia/lung abscess
 - h. Interstitial edema consistent with pneumonia/lung abscess
3. Sterile Site Laboratory Findings
 - i. Any one of the following [i through v]
 - i. Sputum lab findings [any **one** of the following (1, 2)]:
 1. Sputum culture/PCR/serology/gram stain positive for an agent that is not considered a contaminant [see exclusion list below]:
 - a. *Aspergillus* species, *Enterococcus* species, viridians group streptococci, and yeast
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - ii. Blood lab findings [either of the following (1-3)]
 1. Blood culture/PCR/serology positive for an agent that is not considered a contaminant [see exclusion list below]:
 - a. Exclusions
 - i. Coagulase-negative *Staphylococcus*

- ii. *Bacillus spp.* (other than *Bacillus anthracis*)
 - iii. *Corynebacterium spp.*
 - iv. *Propionibacterium spp.*
 - v. *Micrococcus*
 - vi. Diphtheroids
 - vii. Viridians Group Streptococci
 - viii. Enterococci
 - ix. *Clostridium perfringens*
 - x. *Aerococcus*
 - xi. *Alcaligenes faecalis*
 - xii. *Citrobacter*
 - xiii. *Neisseria subflava*
 - xiv. *Stomatococcus*
 - xv. *Streptococcus bovis*
 - xvi. *Veillonella candidemia*
 - xvii. *Mycobacterium tuberculosis*
 - xviii. *S. salivarius*
2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
- iii. Pleural fluid lab findings [either of the following (1, 2)]
 1. Culture/PCR/serology positive for a bacterial pathogen
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - iv. Bronchoscopic specimen or deep endotracheal tube aspiration lab findings [either of the following (1, 2)]
 1. Culture/PCR/serology positive for a bacterial pathogen
 2. Positive viral study (culture/PCR/antigen screen) for a viral pathogen
 - v. Urine antigen detection testing [either of the following (1, 2)]
 1. *Legionella pneumophila*
 2. *Streptococcus pneumoniae*

III. Cellulitis/Soft-Tissue Infection

Both of the following:

1. Any mention of the following with recent onset (≤ 14 days) [*any of the following*]
 - a. Skin erythema
 - b. Surgical site infection
 - c. Superficial central line infection
 - d. Ostomy site infection
 - e. Skin infection with associated lymphangitis
2. Antibiotic treatment initiated for suspected infection

For peer review only

IV. Endocarditis

Any one of the following [1-3]:

1. Major Criteria [both of the following]:
 - a. Suggestive microbiology [at least one of the following]:
 - i. Positive blood culture of an *endocarditis organism* [any of the following]:
 1. *Streptococcus bovis*
 2. *Viridians streptococci*
 3. *Staphylococcus aureus*
 4. *Enterococcus spp.*
 5. HACEK organisms
 6. Coagulase negative staphylococci
 - b. Evidence of endocardial involvement [at least one of the following]:
 - i. New regurgitant murmur (a change in a preexisting murmur does not get scored)
 - ii. Echocardiogram suspicious for any of the following:
 1. Intracardiac mass with no alternative explanation
 2. Endocardial abscess
 3. New partial prosthesis dehiscence
 4. Vegetation on valve
2. Minor Criteria [at least 4 of the following]:
 - a. Predisposing valvular disease or IV drug use
 - b. Temperature $\geq 100.4^{\circ}\text{F}$ or 38°C
 - c. Vascular phenomena
 - i. Janeway lesions, conjunctival hemorrhages, arterial emboli, septic pulmonary infarcts, mycotic aneurysm, intracranial bleed
 - d. Immunologic phenomena
 - i. Osler nodes, Roth Spots, elevated Rheumatoid factor, hematuria in non-catheter urine, or other evidence of glomerulonephritis
 - e. Positive blood cultures
 - i. Excluding a single positive culture for coagulase negative staphylococci or a single positive culture for an organism that does not fall into the “reasonable endocarditis organism” (i.e. coagulase-positive and coagulase-negative *S. aureus*, *Enterococcus*, viridians group Streptococci, *S. bovis*, HACEK organisms)
 - f. Positive serology for Brucella, Bartonella, Legionella, or Chlamydia
 - g. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
3. At least one Major Criteria **AND** 3 minor criteria.

V. Meningitis/Encephalitis

Any one of the following [1 or 2]:

1. Both of the following [a-b]
 - a. Laboratory Findings [any one of the following (i-ix)]
 - i. CSF demonstrates any bacterium
 1. Excluding Diptheroids, Propionibacteria, Bacillus, Coagulase Negative *Staphylococcus*
 - ii. CSF demonstrates Diptheroids, Propionibacteria, Bacillus, Coagulase Negative *Staphylococcus* in the setting of past neurosurgical intervention **AND** physicians elected to treat with antibacterials
 - iii. Blood cultures positive for any of the following:
 1. *S. pneumoniae*
 2. *H. influenza*
 3. *Neisseria meningitidis*
 4. Group B Streptococcus
 - iv. Stool cultures positive for enterovirus
 - v. Throat or sputum cultures positive for *Neisseria meningitidis* in the setting of a rapid onset, overwhelming infection syndrome, including petechiae
 - vi. Serology positive for *Mycoplasma*, *Leptospira*, measles, mumps, lymphocytic choriomeningitis virus, arboviruses (e.g. St. Louis encephalitis virus), or HIV (if historically consistent with acute seroconversion).
 - vii. Brain biopsy demonstrates encephalitis
 - viii. Positive CSF culture or PCR detection for any of the following
 - ix. Acute or convalescent serology demonstrates positive antibody pattern for any of the following:
 1. Encephalitis arbovirus (La Crosse, St. louis, Eastern Equine, Western Equine, Powassan, Japanese, West Nile)
 - b. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected meningitis/encephalitis
2. At least two of the following [a-d]
 - a. Clinical meningitis/encephalitis [at least two of the following]:
 - i. Petechial rash
 - ii. Nuchal rigidity (by history or exam)
 - iii. Altered sensorium
 - iv. Fever
 - v. Altered level of consciousness, including “agitation” or “lethargy”
 - vi. Behavioral change
 - vii. Diminished level of consciousness (not easily roused)
 - viii. History of any of the following: headaches, altered mental status, or recent exposure to patient with known bacterial meningitis
 - ix. Reduction in fever within 72 hours of starting anti-bacterial
 - b. Inflammatory CSF [at least one of the following i-ii]
 - i. Pleocytosis: ≥ 15 WBC/mm³ (after subtracting one WBC for every 1,000 RBC)
 - ii. Elevated protein (based on local lab-determined upper limits)
 - c. Suggestive Findings [at least one of the following (i-iv)]

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
- i. Septic syndrome
 - ii. Focal neurological deficits documented during examination (such as flaccid paralysis or speech alterations for West Nile Virus)
 - iii. Abnormal imaging
 1. Computed tomography or magnetic resonance imaging (MRI) demonstrating focal edema or inflammation or hemorrhage
 2. Indicated as “meningitis/encephalitis” or “compatible with meningitis/encephalitis” or “cannot rule out meningitis/encephalitis”
 - iv. Findings indicating an abnormal electroencephalography (such as focal periodic discharges)
- d. Antibiotic/antiviral/antifungal treatment initiated/recommended for presumed meningitis/encephalitis

VI. Pyelonephritis

At least two of the following [1-4]:

1. Suggestion of infection [at least one of the following]:
 - a. Temperature \geq 100.4°F (38°C)
 - b. Peripheral blood WBC \geq 10,000/mm³
 - c. Positive blood culture for any of the following:
 - i. Gram Negative Rods
 - ii. *Enterococcus spp.*
 - iii. *Staphylococcus saprophyticus*
 - d. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
2. Strong renal localization [at least one of the following]:
 - a. CT, MRI, or Ultrasound Suggestive of Renal Inflammation
3. Minor Criteria [at least two of the following]:
 - a. Flank pain
 - b. Costovertebral angle tenderness
 - c. Complaints of dysuria, frequency, or suprapubic pain
 - d. Any pyuria
 - e. Urine culture positive for a single organism
4. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected pyelonephritis

VII. Septic Arthritis/Osteomyelitis

Any one of the following (1-5):

1. Synovial fluid gram stain or tissue gram stain or special stain demonstrating any organism
2. Joint culture/PCR/serology positive for any organism
3. At least two of the following (a-d):
 - a. Positive blood culture/PCR/serology
 - b. Joint with acute (≤ 7 days) worsening of inflammatory features (**at least two of the following**):
 - i. Pain on history
 - ii. ROM
 - iii. Warmth
 - iv. Effusion
 - v. Swelling
 - vi. Limited range of motion
 - c. Antibiotic/antiviral/antifungal/antifungal treatment initiated/recommended for suspected infection
 - d. Any one of the following (i-iv)
 - i. Synovial fluid WBC $\geq 30,000/\text{mm}^3$
 - ii. Synovial fluid WBC $\geq 60,000/\text{mm}^3$ with $> 75\%$ PMNs
 - iii. Skin lesions, tenosynovitis, or urethral/cervical/rectal Gram stain or culture suggestive of *Neisseria gonorrhoeae*
 - iv. Any indication of the following in the synovial fluid: needle-like crystals, CPPD crystals, uric acid.
4. Positive bone biopsy [at least one of the following (a-c)]:
 - a. Positive culture for any organism
 - b. Positive gram stain
5. Imaging and indirect features [**at least two of the following (a-c)**]:
 - a. Consistent imaging [at least one of the following (i-iv)]:
 - i. Plain X-ray read by a radiologist as suggestive of osteomyelitis
 - ii. CT Scan read by a radiologist as suggestive of osteomyelitis
 - iii. MRI read by a radiologist as suggestive of osteomyelitis
 - iv. Bone scan or WBC scan read as suggestive of osteomyelitis
 - b. Suggestive indirect features[at least one of the following (i-viii)]:
 - i. Temperature $> 100.4^\circ\text{F}$ (38°C)
 - ii. Bony pain or tenderness or erythema over bone suspected to be infected
 - iii. Draining soft tissue sinus over bone suspected to be infected
 - iv. Positive “probe to bone” (or visible bone in deep ulcer at suspected site)
 - v. Blood culture positive for *S. aureus*
 - vi. ESR ≥ 75 mm/hour
 - vii. Intravenous drug use or indwelling catheter
 - viii. Inflammation on imaging associated with an orthopedic prosthesis
 - c. Positive culture for any organism from wound sample over the bone suspected of infection
 - d. Antibiotic/antiviral/antifungal treatment for suspected infection

VIII. References

1. Grijalva CG, Chung CP, Stein CM, et al. Computerized definitions showed high positive predictive values for identifying hospitalizations for congestive heart failure and selected infections in Medicaid enrollees with rheumatoid arthritis. *Pharmacoepidemiology and drug safety* 2008; 17(9): 890-5.
2. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital discharge databases coded serious bacterial infections accurately. *Journal of clinical epidemiology* 2007; 60(4): 397-409.
3. Patkar NM, Curtis JR, Teng GG, et al. Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients. *Journal of clinical epidemiology* 2009; 62(3): 321-7, 7.e1-7.

Section & Topic	No	Item	Reported on page #
Title or Abstract			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
<i>Participants</i>	6	Eligibility criteria	4,5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4,5
	9	Whether participants formed a consecutive, random or convenience series	4-6
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5, Table 1, Supplementary Appendix
	10b	Reference standard, in sufficient detail to allow replication	6, Supplementary Appendix
	11	Rationale for choosing the reference standard (if alternatives exist)	6, Supplementary Appendix
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6, Supplementary Appendix
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6, Supplementary Appendix
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6, Supplementary Appendix
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	7,8
	15	How indeterminate index test or reference standard results were handled	7,8
	16	How missing data on the index test and reference standard were handled	7,8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7,8
	18	Intended sample size and how it was determined	7,8
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	8
	20	Baseline demographic and clinical characteristics of participants	8
	21a	Distribution of severity of disease in those with the target condition	n/a
	21b	Distribution of alternative diagnoses in those without the target condition	n/a
	22	Time interval and any clinical interventions between index test and reference standard	n/a
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	9,10, Table 2
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	9,10 Table 2
	25	Any adverse events from performing the index test or the reference standard	n/a
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	11-13
	27	Implications for practice, including the intended use and clinical role of the index test	11-13

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

OTHER INFORMATION			
	28	Registration number and name of registry	n/a
	29	Where the full study protocol can be accessed	14
	30	Sources of funding and other support; role of funders	13

For peer review only

BMJ Open: first published as 10.1136/bmjopen-2017-020857 on 19 June 2018. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.