OPEN

Accuracy of *International Classification of Diseases*, 10th Revision Codes for Identifying Sepsis: A Systematic Review and Meta-Analysis

OBJECTIVES: Administrative databases are increasingly used in research studies to capture clinical outcomes such as sepsis. This systematic review and meta-analysis examines the accuracy of *International Classification of Diseases*, 10th revision (ICD-10), codes for identifying sepsis in adult and pediatric patients.

DATA SOURCES: We searched MEDLINE, EMBASE, Web of Science, CENTRAL, Epistemonikos, and McMaster Superfilters from inception to September 7, 2021.

STUDY SELECTION: We included studies that validated the accuracy of sepsis ICD-10 codes against any reference standard.

DATA EXTRACTION: Three authors, working in duplicate, independently extracted data. We conducted meta-analysis using a random effects model to pool sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). We evaluated individual study risk of bias using the Quality Assessment of Diagnostic Accuracy Studies tool and assessed certainty in pooled diagnostic effect measures using the Grading of Recommendations Assessment, Development, and Evaluation framework.

DATA SYNTHESIS: Thirteen eligible studies were included in the qualitative synthesis and the meta-analysis. Eleven studies used manual chart review as the reference standard, and four studies used registry databases. Only one study evaluated pediatric patients exclusively. Compared with the reference standard of detailed chart review and/or registry databases, the pooled sensitivity for sepsis ICD-10 codes was 35% (95% CI, 22–48, low certainty), whereas the pooled specificity was 98% (95% CI: 98–99, low certainty). The PPV for ICD-10 codes ranged from 9.8% to 100% (median, 72.0%; interquartile range [IQR], 50.0–84.7%). NPV ranged from 54.7% to 99.1% (median, 95.9%; interquartile range, 85.5–98.3%).

CONCLUSIONS: Sepsis is undercoded in administrative databases. Future research is needed to explore if greater consistency in ICD-10 code definitions and enhanced quality measures for ICD-10 coders can improve the coding accuracy of sepsis in large databases.

KEY WORDS: administrative; *International Classification of Diseases*, 10th revision codes, medical records; sepsis; systematic review; validation

Bonnie Liu, BHSc[1]

Milena Hadzi-Tosev, MSc[1]

Yang Liu, MMath[1]

Kayla J. Lucier, MPH[1]

Anchit Garg, BHSc[1]

Sophie Li, BHSc[1]

Nancy M. Heddle, MSc[1,2,3]

Bram Rochwerg, MD, MSc[2,4]

Shuoyan Ning, MD[1,2,5,6]

Sepsis is characterized by life-threatening organ dysfunction caused by dysregulated host responses to infection (1, 2). In 2017, sepsis was estimated to affect 48.9 million people globally and caused 11.0 million deaths, representing 19.7% of all global deaths (3).

KEY POINTS

- **Question:** The aim of this study is to examine the accuracy of *International Classification of Diseases*, 10th revision (ICD-10), codes when used to identify sepsis among pediatric and adult patients.
- **Findings:** Our systematic review and meta-analysis found that compared with reference standards, the pooled sensitivity and specificity of ICD-10 sepsis codes were 35% (95% CI, 22–48%, low certainty) and 98% (95% CI, 98–99%, low certainty), respectively.
- **Meaning:** Overall, sepsis is undercoded by ICD-10 codes in administrative databases and subject to high heterogeneity.

The World Health Assembly has called for improved research, monitoring, prevention, diagnosis, and management of sepsis globally (4). The *International Classification of Diseases* (ICD), a standardized set of codes created by the World Health Organization, are relied upon to inform accurate epidemiologic monitoring, global mortality estimates, insurance reimbursement, resource allocation, and database-driven research of sepsis (5, 6). At most institutions, ICD codes are manually applied by trained health record coders to capture diseases and injuries over the course of a patient's hospital stay (7). After undergoing many revisions for improvements in its use and structure, the 10th revision of ICD (ICD-10) was endorsed by the World Health Organization in 1990 and is used in over 100 countries (8). The quality of ICD-10 coding is influenced by not only the clarity, precision, and completeness of patient charts but also the accuracy and consistency of health record coders (9).

Epidemiologic trends suggest that although the incidence of sepsis is increasing, sepsis-associated mortality is decreasing (10–15). However, other studies have found that the age-standardized incidence of sepsis is decreasing (3) and that the combined rate of death or discharge to hospice in sepsis patients has remained stable (15). Although temporal trends and variations in care may explain these discrepancies, it is also possible that coding disparities, claims-based coding bias, and other factors may influence these conflicting trends (16).

ICD-10 codes are increasingly used to identify sepsis in administrative databases to capture this important outcome, but their accuracy for the identification of sepsis remains unclear. During the transition of 9th revision of ICD (ICD-9) to ICD-10, several studies that broadly compared the accuracy of ICD-9 to ICD-10 codes reported comparable sensitivity values for selected conditions (17, 18). For sepsis, a 2007 study that directly compared ICD-9 and ICD-10 reported higher sensitivity for sepsis and lower sensitivity for community-acquired sepsis for ICD-10 codes; ICD-9 had higher specificity for sepsis but lower specificity for community-acquired sepsis when compared with ICD-10 codes (19). Additionally, the authors reported that ICD-10 code sensitivity for sepsis and community-acquired sepsis differed significantly, whereas ICD-9 code sensitivity for the two diagnoses was similar. To date, systematic reviews and meta-analyses of ICD validation for sepsis have focused on ICD-9 or have pooled ICD-9 with ICD-10 rather than reporting ICD-10 alone (20, 21). Since many international health systems have replaced ICD-9 use with ICD-10, ICD-10-specific sepsis validation is necessary for understanding the current quality of sepsis code reporting (22).

This has implications for academic research on sepsis and health insurance reimbursement (23). The aim of this systematic review and meta-analysis is to examine the accuracy of ICD-10 codes when used to identify sepsis among pediatric and adult patients.

MATERIALS AND METHODS

Data Sources and Searches

We developed this systematic review and meta-analysis to conform to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) standards, subset Diagnostic Test Accuracy (DTA) (24). The PRISMA-DTA checklist is provided in **Supplementary Table 1** (http://links.lww.com/CCX/B81). The protocol was not registered.

We searched the following electronic databases from inception until September 7, 2021: Medical Literature Analysis and Retrieval System Online (MEDLINE) (1946–), Excerpta Medica Database (EMBASE) (1974–), Institute for Scientific Information (ISI) Web of Science (1970–), Cochrane Central Register of Controlled Trials (CENTRAL), Epistemonikos, and McMaster SuperFilters. The search strategy consisted of: 1) terms

to identify ICD-10 codes; 2) keywords and Medical Subject Heading terms relating to sources of codes, validation study types, and measures of statistical accuracy; and 3) terms to identify sepsis or septic shock. The full search strategies are provided in the **Supplementary Appendix 1** (http://links.lww.com/CCX/B81).

We did not apply any restrictions for date, publication status, study design (e.g., retrospective, prospective, randomized, case cohort, systematic review), or case definition, although only English language studies were included. If the search strategy identified a systematic review, references were screened for additional eligible studies.

Study Selection

Potential citations were screened in two stages. First, three investigators (B.L., A.G., S.L.) independently and in duplicate screened the titles and abstracts against the inclusion criteria and selected studies for full-text review. In the second stage, the same reviewers screened the full-texts in duplicate for final study inclusion. Disagreements were resolved through discussion and third-party adjudication (K.J.L., S.N.).

We included studies that met the following criteria: 1) evaluated the accuracy of ICD-10 codes for sepsis; 2) applied any reference standard such as chart review, blood cultures, or other patient registries; and 3) reported at least one measure of accuracy: sensitivity, specificity, positive predictive value (PPV), or negative predictive value (NPV) or provided the information necessary to craft 2×2 tables for calculation. We are choosing to use the term reference standard instead of gold standard because all reference standards used by studies may not have been the best standard available to investigators.

We excluded studies if they: 1) combined both ICD-9 and ICD-10 in their analysis, 2) lacked a clear reference standard, or 3) used ICD-9 as reference standard. There were no restrictions for age, sex, region, or clinical setting. We did not require studies to provide the full list of ICD-10 codes used.

Data Extraction and Quality Assessment

We developed a data abstraction form and collected the following information from the final included studies: article title, first author name, year of publication, country of origin, study duration, method of identifying

sepsis, patient population (inpatient or outpatient), patient age (adult or pediatric), and any patients grouped by their diagnoses. Additionally, we collected the study's definition of sepsis, validation procedure and reference standard, sample size, outcomes (e.g., sensitivity, specificity, PPV, NPV), and any subgroup analyses. ICD-10 codes used and their positions (e.g., primary, secondary, or any position) were documented if provided. Data were abstracted independently and in duplicate (B.L., A.G., S.L.) and compared for consistency. Disagreements were resolved through discussion or third-party adjudication if necessary (K.J.L., S.N.).

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool was used to assess the risk-of-bias of included studies (25). All studies were reviewed independently by at least two investigators (B.L., A.G., S.L.) trained in applying the QUADAS tool, and any discordance was resolved through discussion and consensus. To evaluate the certainty of evidence in the pooled diagnostic characteristics, two investigators (B.L., A.G.) independently assessed the pooled evidence using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach (26).

Data Analysis

We evaluated the PPV, NPV, sensitivity, and specificity for sepsis ICD-10 codes compared with the reference standard. We performed meta-analysis for sensitivity and specificity using a DerSimonian-Laird (27) and Jackson (28) method random effect model and inverse variance method for study weights. If SE was not reported, we used the 95% CIs to calculate SE to allow for meta-analysis using the formula: $\theta \pm 1.96 * s.e(\theta)$. We assumed a Wald's CI method, unless otherwise stated.

Three secondary analyses were performed. First, we compared the diagnostic accuracy of ICD-10 codes with manual chart review as reference standard compared with other reference standards. Second, we compared the accuracy of codes when studies provided a full list of codes compared with studies that did not. Finally, to determine whether certain ICD-10 codes had lower levels of accuracy at capturing sepsis diagnosis, we evaluated studies that provided individual codes across all PPV, NPV, sensitivity, and specificity and categorized them into quartiles of accuracy. We chose the 25% quartile as a cut-off point for lower accuracy values and the 75% quartile as a cut-off point for higher accuracy values.

Two analyses were performed to verify the accuracy of our sensitivity and specificity meta-analyzed estimates given the possible intercorrelation and interdependence between values from the same patient cohort/studies. In the first analyses, we selected only one subgroup from each study for inclusion in the random effects meta-analysis. Details on the subgroup selection process can be found in **Supplementary Appendix 2** (http://links.lww.com/CCX/B81). In the second analyses, we first pooled the subgroups of each study to generate study-specific sensitivity and specificity values; we then pooled these study-specific values for a final sensitivity and specificity value.

All analyses were performed using R Version 4.0.2 (29) with meta package Version 4.16-2 (R Foundation for Statistical Computing, Vienna, Austria). We assessed statistical heterogeneity using the I^2 statistic, the chi-square test (30), and visual inspection of the forest plots.

RESULTS

Study Characteristics

After removing duplicates, we identified a total of 878 citations from our search, of which 42 underwent full text review, and ultimately 13 eligible studies were included in the systematic review and meta-analysis (**Fig. 1**). The results of individual sepsis studies are summarized in **Supplementary Table 2** (http://links.lww.com/CCX/B81). **Supplementary Table 3** (http://links.lww.com/CCX/B81) summarizes each study's methodology.

The 13 included studies were published from 1998 to 2020, with data collection between 1994 and 2018. Sample size ranged from 78 to 1,645 participants. Four studies were conducted in Denmark (31–34), three in Australia (35–37), one in Germany (38), two in Canada (39, 40), one in England (41), one in Switzerland (42), and one in the United States (43).

Eleven studies were retrospective observational studies, and two were prospective observational studies (42, 43). ICD-10 codes were evaluated in insurance claims data (38), national patient registries (31, 32, 34), hospital discharge registries/data (33, 36, 39, 40, 42), emergency department information systems and hospital morbidity data system (35), perioperative registries (37), and National Health Service hospital episode statistics (41).

Septic shock was included by three studies (37, 38, 41). Severe sepsis was included by three studies and

was defined as the ICD-10 code R65.1! (38), an Acute Physiology and Chronic Health Evaluation II diagnosis with or without additional disease ICD-9 codes (35), and a combination of sepsis codes and at least one organ dysfunction code (40). Two studies, which used bacteremia database and blood cultures, included septicemia (31, 33). The remaining studies included sepsis, with severity not explicitly outlined.

In terms of patient population, one study evaluated pediatric patients exclusively (42); one evaluated both pediatric and adult in-patients (37); seven exclusively evaluated adult in-patients (31, 32, 34, 35, 38, 40, 43); one evaluated adult outpatients in combination with inpatients, accident, emergency, and surgery patients (41); one evaluated postoperative patients (age unspecified) (39); and two evaluated hospital patients (age and patient type unspecified) (33, 36).

Of the 13 studies reviewed, seven used manual chart review exclusively as the reference standard for sepsis diagnosis (32, 34, 37–40, 43). One study used manual chart review paired with a bacteremia database as the reference standard (33), whereas three studies used laboratory information systems/blood cultures (31, 36, 42). Additionally, one study used an ICU clinical database (35), where sepsis was confirmed by an Acute Physiology Chronic Health Evaluation score and up to four corresponding ICD-9-CM codes, and one used a National Surgical Quality Improvement database (41), where sepsis was confirmed by matching text strings.

Eleven studies reported their ICD-10 sepsis codes (31–41), whereas two studies did not report their codes (42, 43). Full list of ICD-10 codes used by studies is provided in Supplementary Table 3 (http://links.lww.com/CCX/B81). Sepsis diagnosis was in the primary position for two studies (34, 35), in the "any" position for three studies (16, 32, 40), in the secondary position for one study (31), and in an unknown position for the remaining studies. Overall, codes in the primary position had comparable PPVs (68.9–93.9%) to codes in "any" position (50–100%) (16, 32, 34, 35, 40). However, one study (Søgaard et al [31]) found that sepsis codes in the secondary position had a higher PPV than codes in the primary or "any" position.

Risk of Bias Assessment

Supplementary Figure 1 (http://links.lww.com/CCX/B81) summarizes risk of bias assessments using
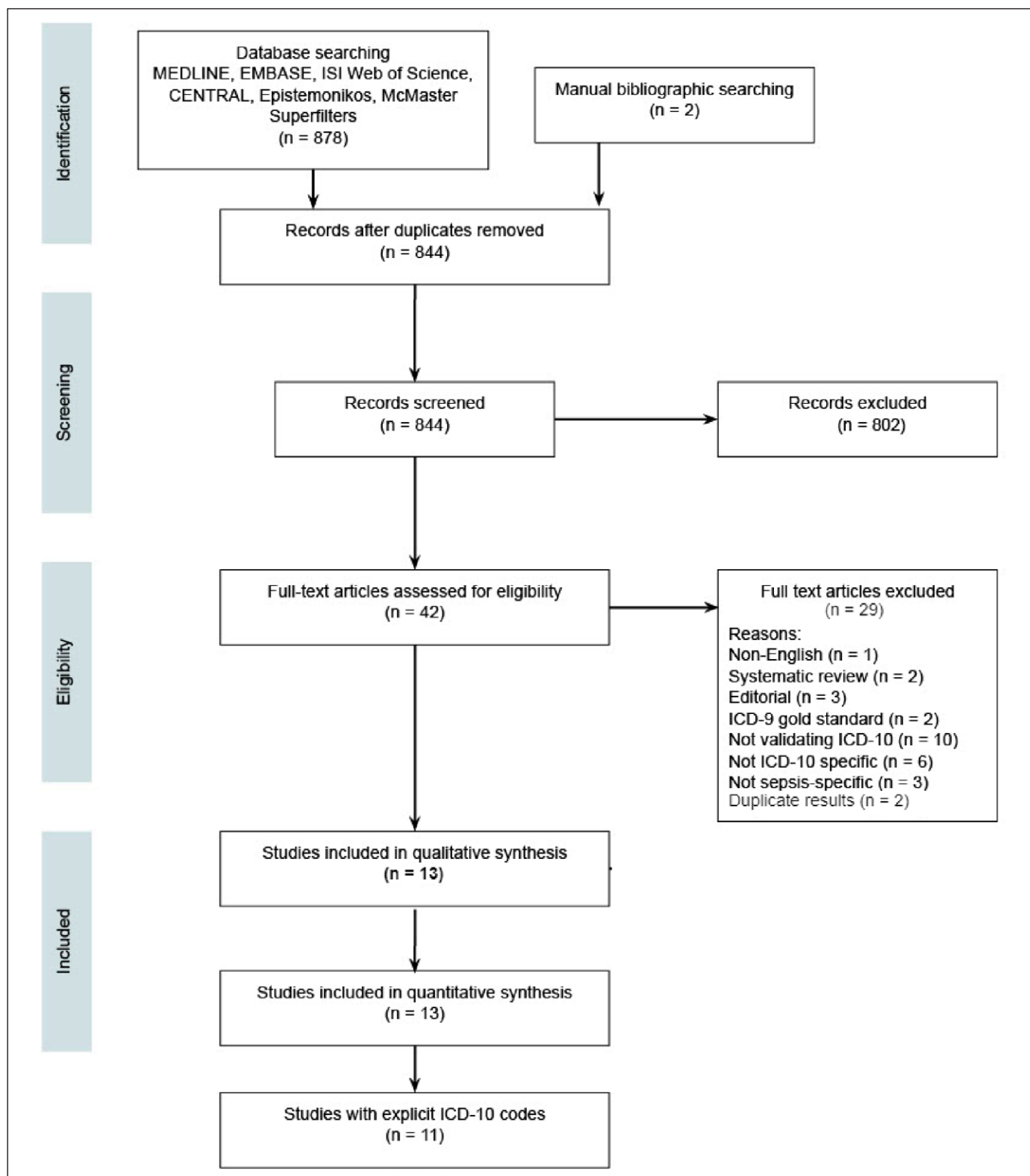
Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. PRISMA flow diagram showing studies identified, included and excluded, and reasons for exclusion during full text review. MEDLINE = Medical Literature Analysis and Retrieval System Online; EMBASE = Excerpta Medica Database; ISI = Institute of Scientific Information.

QUADAS. Most articles had a low risk of bias, resulting from incomplete reporting of missing data, vague index test and reference standard descriptions, and the investigator's awareness of a chart's ICD-10 codes when completing the sepsis assessment as a reference standard.

Synthesis of Results

Supplementary Table 4 (http://links.lww.com/CCX/B81) shows the nonpooled and unweighted mean, median, sample SD, and range of sepsis ICD-10 code PPV, NPV, sensitivity, and specificity.

Sensitivity was provided by nine studies (33, 35–38, 40–43) and ranged from 2.2% to 71.9% (median, 41.9%; interquartile range [IQR], 19.3–57.5%). Six studies (35, 37, 38, 40, 41, 43) provided specificity, which ranged from 85.4% to 100% (median, 99.5%; IQR, 96.2–99.6%). Twelve studies (31–41, 43) reported PPV, which ranged from 9.8% to 100% (median, 72%; IQR, 50–84.7%). Six studies (35, 37, 38, 40, 41, 43) reported NPV, which ranged from 54.7% to 99.1% (median, 95.9%; IQR, 85.5–98.3%).

Compared with a reference standard, pooled sensitivity was 35% (95% CI, 22–48%; 5 studies; low certainty) (**Fig. 2A**), whereas the pooled specificity was 98% (95% CI, 98–99%; 3 studies; low certainty) (**Fig. 2B**). Given that studies reported different sensitivity and specificity estimates based on population type (e.g., ICU, non-ICU), abstraction method (e.g., implicit coding which abstracts sepsis cases based on infectious disease and organ dysfunction codes so as to mirror the clinical sepsis criteria, explicit coding which abstracts sepsis cases based on all sepsis codes, or R codes which are clinical sepsis codes related to general signs and symptoms) (38), administrative database source (e.g., emergency department information system, hospital mortality data system), and sepsis type (e.g. sepsis, severe sepsis, septicemia), we isolated all subgroups for the sensitivity and specificity analysis.

Supplementary Appendix 2 (http://links.lww.com/CCX/B81) details our two analyses, which reassessed and confirmed the accuracy of our pooled sensitivity and specificity estimates. The first analysis, which used one isolated subgroup from each study, produced a pooled sensitivity of 33% (95% CI, 5–61%) and pooled specificity of 97% (95% CI, 95–99%). The second analysis, which pooled subgroups and then pooled conglomerate study values, produced a sensitivity of 32% (95% CI, 6–58%) and pooled specificity of 98% (95% CI, 97–99%).

Supplementary Table 5 (http://links.lww.com/CCX/B81) summarizes the GRADE evaluation for the accuracy of sepsis ICD-10 codes. Certainty of evidence was rated down by two levels for inconsistency for both sensitivity and specificity. There was significant clinical and methodological heterogeneity (e.g., use of different ICD-10 codes, outcome definitions, reference standards), which also translated into significant statistical heterogeneity (I^2 of pooled sensitivity = 99%; I^2 of pooled specificity = 94%). The five studies included in the sensitivity analysis and the three studies included in the specificity analysis did not demonstrate significant risk of bias because only one study in both analyses failed to blind investigators to patient's ICD-10 codes. It is important to note that the low risk of bias for the sensitivity and specificity analyses is separate from the QUADAS evaluation, which evaluated all 13 studies cumulatively. We have incorporated and contextualized these concerns around inconsistency into the low certainty of evidence for sensitivity and specificity using GRADE evaluation.

Our secondary analyses showed that studies exclusively using manual chart review as reference standard yielded a lower median PPV but higher median sensitivity compared with studies that did not use manual chart review exclusively (**Supplementary Table 6**, http://links.lww.com/CCX/B81). Secondary analyses demonstrated that studies that explicitly reported their specific ICD-10 codes had a lower mean PPV, NPV, and sensitivity compared with studies that did not report their codes (**Supplementary Table 7**, http://links.lww.com/CCX/B81); however, this finding may be confounded by the greater number of studies, and thus heterogeneity, reporting explicit codes. Of the 303 different codes captured in this systematic review, the use of code A41.9A (sepsis with unspecified organism) corresponded to the highest quartile of diagnostic accuracy as it exceeded the 75% quartile of PPV, NPV, sensitivity, and specificity. However, codes O85.9, P36.0, P36.1, P36.2, P36.3, P36.4, P36.5, P36.8, P36.9, and P36 corresponded to the lowest quartile of diagnostic accuracy (**Supplementary Table 8**, http://links.lww.com/CCX/B81).

DISCUSSION

This systematic review and meta-analysis evaluated the accuracy of ICD-10 codes for identifying sepsis in pediatric and adult patients. We found that sepsis ICD-10 codes have consistently high NPV (median, 95.9%, IQR, 85.5–98.3%) and specificity (median, 99.5%; IQR,
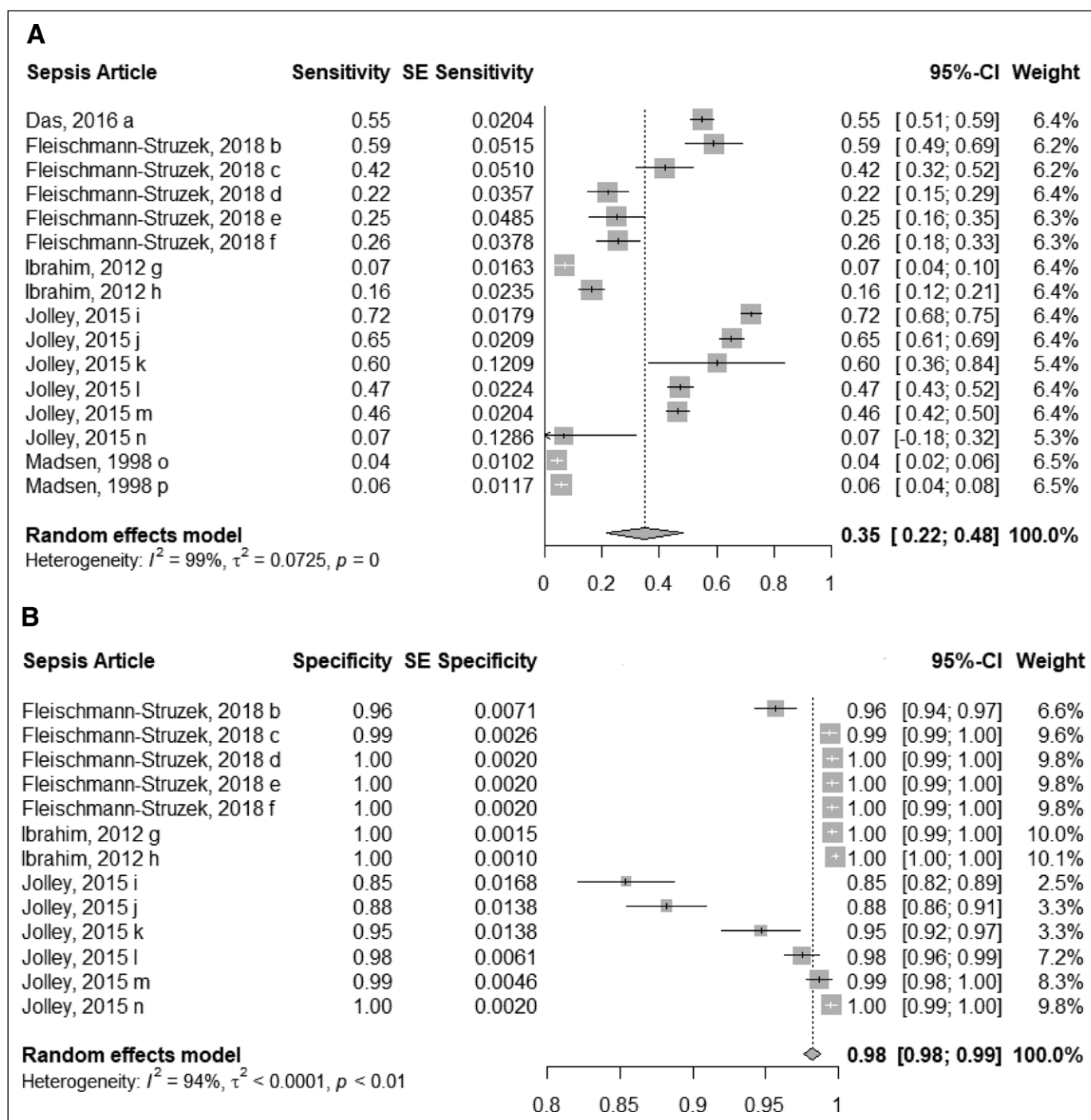
A

Sepsis Article	Sensitivity	SE Sensitivity		95%-CI	Weight
Das, 2016 a	0.55	0.0204		0.55 [0.51; 0.59]	6.4%
Fleischmann-Struzek, 2018 b	0.59	0.0515		0.59 [0.49; 0.69]	6.2%
Fleischmann-Struzek, 2018 c	0.42	0.0510		0.42 [0.32; 0.52]	6.2%
Fleischmann-Struzek, 2018 d	0.22	0.0357		0.22 [0.15; 0.29]	6.4%
Fleischmann-Struzek, 2018 e	0.25	0.0485		0.25 [0.16; 0.35]	6.3%
Fleischmann-Struzek, 2018 f	0.26	0.0378		0.26 [0.18; 0.33]	6.3%
Ibrahim, 2012 g	0.07	0.0163		0.07 [0.04; 0.10]	6.4%
Ibrahim, 2012 h	0.16	0.0235		0.16 [0.12; 0.21]	6.4%
Jolley, 2015 i	0.72	0.0179		0.72 [0.68; 0.75]	6.4%
Jolley, 2015 j	0.65	0.0209		0.65 [0.61; 0.69]	6.4%
Jolley, 2015 k	0.60	0.1209		0.60 [0.36; 0.84]	5.4%
Jolley, 2015 l	0.47	0.0224		0.47 [0.43; 0.52]	6.4%
Jolley, 2015 m	0.46	0.0204		0.46 [0.42; 0.50]	6.4%
Jolley, 2015 n	0.07	0.1286		0.07 [-0.18; 0.32]	5.3%
Madsen, 1998 o	0.04	0.0102		0.04 [0.02; 0.06]	6.5%
Madsen, 1998 p	0.06	0.0117		0.06 [0.04; 0.08]	6.5%
Random effects model				**0.35 [0.22; 0.48] 100.0%**	

Heterogeneity: $I^2 = 99\%$, $\tau^2 = 0.0725$, $p = 0$

B

Sepsis Article	Specificity	SE Specificity		95%-CI	Weight
Fleischmann-Struzek, 2018 b	0.96	0.0071		0.96 [0.94; 0.97]	6.6%
Fleischmann-Struzek, 2018 c	0.99	0.0026		0.99 [0.99; 1.00]	9.6%
Fleischmann-Struzek, 2018 d	1.00	0.0020		1.00 [0.99; 1.00]	9.8%
Fleischmann-Struzek, 2018 e	1.00	0.0020		1.00 [0.99; 1.00]	9.8%
Fleischmann-Struzek, 2018 f	1.00	0.0020		1.00 [0.99; 1.00]	9.8%
Ibrahim, 2012 g	1.00	0.0015		1.00 [0.99; 1.00]	10.0%
Ibrahim, 2012 h	1.00	0.0010		1.00 [1.00; 1.00]	10.1%
Jolley, 2015 i	0.85	0.0168		0.85 [0.82; 0.89]	2.5%
Jolley, 2015 j	0.88	0.0138		0.88 [0.86; 0.91]	3.3%
Jolley, 2015 k	0.95	0.0138		0.95 [0.92; 0.97]	3.3%
Jolley, 2015 l	0.98	0.0061		0.98 [0.96; 0.99]	7.2%
Jolley, 2015 m	0.99	0.0046		0.99 [0.98; 1.00]	8.3%
Jolley, 2015 n	1.00	0.0020		1.00 [0.99; 1.00]	9.8%
Random effects model				**0.98 [0.98; 0.99] 100.0%**	

Heterogeneity: $I^2 = 94\%$, $\tau^2 < 0.0001$, $p < 0.01$

Figure 2. Sensitivity and specificity forest plots of sepsis *International Classification of Diseases*, 10th revision (ICD-10) codes. Forest plots of the sensitivity (5 studies; **A**) and specificity (3 studies; **B**) of sepsis ICD-10 codes compared with the reference standard of detailed chart review and/or registry databases. a = A41.0, b = implicit abstraction for severe sepsis, c = explicit abstraction with organ dysfunction codes for severe sepsis, d = R code (R65.0, R65.1, R57.2) abstraction for sepsis, e = R code (R65.1, R57.2) abstraction for severe sepsis, f = explicit abstraction for sepsis, g = diagnosis-based code categories in emergency department information system, h = diagnosis-based code categories in hospital mortality data system, i = optimized coding algorithm for ICU sepsis, j = optimized coding algorithm for ICU severe sepsis, k = optimized coding algorithm for non-ICU sepsis, l = CIHI coding algorithm for ICU severe sepsis, m = CIHI coding algorithm for ICU sepsis, *n* = Canadian Institute for Health Information (CIHI) coding algorithm for non-ICU Sepsis, o = septicemia, p = sepsis and septicemia.

96.2–99.6%), but low sensitivity (median, 41.9%, IQR, 19.3–57.5%) and low PPV (median, 72.0%; IQR, 50.0–84.7%) compared with chart review and registries. The

low sensitivity indicates that patients with sepsis are often not coded with sepsis ICD-10 codes, whereas the high specificity indicates the accuracy of such ICD-10

codes when present. Thus, our findings highlight an issue of undercoding. We also found that there is significant heterogeneity in the sensitivity (I^2 = 99%) and specificity (I^2 = 94%) of sepsis ICD-10 codes in clinical, claims, and administrative or registry databases. Overall, the certainty of evidence is rated as low.

These findings are consistent with published literature on the coding accuracy of ICD-9 and ICD-10 sepsis codes. Previous systematic reviews from 2013 and 2015 reported that ICD-9 and ICD-10 coding failed to comprehensively capture sepsis (sensitivity median, 42.4%) and that both coding systems demonstrated high heterogeneity across studies (sensitivity range, 5.9–82.3%) (20, 21). It should be noted however that ICD-10 codes in previous studies were not independently assessed from ICD-9 codes. In our systematic review, the sensitivity of ICD-10 codes for severe sepsis and septic shock ranged 7.1–65.1% and 2.22–59.0%, respectively; this range is comparable with the range of sensitivity of ICD-9 codes for severe sepsis (20.5–50.3%) and septic shock (42.4–75.1%) (20). Previous studies have highlighted the underreporting of sepsis (19). One population-based survey found that the incidence of severe sepsis was 7.1 times lower when ICD-10 codes were used compared with when symptom and clinical findings were used to identify sepsis (44). Our systematic review and meta-analysis provide a timely confirmatory update to the literature, as ICD-9 codes become progressively outdated.

A number of factors have been identified which may play a role in sepsis undercoding. A 2015 study showed that sepsis codes in non-ICU patients had a lower sensitivity compared with ICU patients, even when coding strategies were optimized (40); the authors hypothesized that sepsis prevalence in non-ICU settings may be playing a role because non-ICU physicians, who may encounter sepsis less frequently due to lower disease prevalence, may identify, and document sepsis differently than ICU physicians. Although depending on the institution and clinical practice variations, prevalence of sepsis may vary and may be significantly higher in non-ICU settings at select centers (e.g., emergency departments) (45, 46). Other factors such as surgical status and timing of disease onset may also be important for sepsis identification; one study reported that patients presenting with postoperative sepsis at admission had ICD-10 codes with lower PPV compared with patients who developed postoperative sepsis during hospital stay (39). The sensitivity of ICD-10 coding is also affected by the type of sepsis evaluated (general vs community-acquired) (19), the site of infection (47), and the severity of sepsis (48). In our study, we identified that the use of code A41.9A (sepsis with unspecified organism) corresponded to the highest quartile of diagnostic accuracy, although sensitivity was still limited at 57.5%. The importance of coding abstraction strategies (35, 38) and the human challenges that come with coding (clerical errors, experience and training differences, challenges with the interpretation of healthcare documentation) cannot be understated. The qualifications, training, and years of experience of medical coders may influence the validity of ICD-10 coding, leading to variability between included studies (49).

Considerations to reduce undercoding may include having specialized coding procedures for ICU patients (where sepsis is most common) (50), including sepsis codes found in primary and secondary diagnostic fields when capturing outcomes (51) and pairing sepsis ICD-10 codes and electronic medical record data to other data sources (e.g., laboratory, pharmacy, microbiology data) or with other clinical factors (e.g., heart rate, respiratory rate, body temperature, WBC count, and markers for organ dysfunction) (19, 20). Combination coding algorithms are another consideration, although our review found mixed results on their efficacy. One study (Fleischmann-Struzek et al [38]) reported that adding organ dysfunction codes increased the ability of explicit sepsis codes to rule out disease (sensitivity) but decreased the probability that patients with those codes truly had sepsis (PPV); another study (Jolley et al [40]) found that appending an optimized set of sepsis codes to existing CIHI code definitions increased sensitivity and NPV but decreased specificity and PPV. Finally, one study (Lauridsen et al [32]) confirmed that adding therapy codes, specifically inotropic and vasopressor codes, substantially increased the probability that those with septic shock codes had experienced septic shock (PPV). The use of an implicit sepsis coding strategy, which abstracts sepsis cases based on the combination of infection codes and organ dysfunction codes rather than sepsis documentation, has been investigated by two studies in our systematic review (35, 38). An implicit coding strategy increased the sensitivity of severe sepsis ICD-10 codes (R65.1!, R57.2) by 33.9% and sepsis ICD-10 codes by 72.9% at a cost of PPV decrease in 34% and 44.8%, respectively (38).

There is a need for new studies which test novel ways of applying sepsis diagnostic codes, methods of data linkage, and combination coding algorithms to improve the diagnostic accuracy of ICD-10 codes for sepsis. Since coding can be confounded by changing policies, other efforts such as clinical documentation improvement programs, continuous coder training and audits, and diagnosis cover sheets or financial incentives may also be considered (52–55).

Ultimately, the undercoding of sepsis ICD-10 codes in clinical, claims, and administrative/registry databases adversely impacts epidemiologic monitoring, global mortality estimates, insurance reimbursement, resource allocation, and database-driven research of sepsis (40).

Since ICD codes are commonly used as outcome definitions, the inaccuracy of sepsis ICD-10 codes may also have implications on the development of machine-learning algorithms and electronic predictive tools for early identification of sepsis (56), as well as the implementation of big data analytics in healthcare (57). Underdiagnosis of sepsis by physicians may exacerbate the undercoding of sepsis ICD-10 codes since coders rely on physician documentation for coding. In one study (Rezende et al [58]), only 31% of septic patients seen in the emergency department were diagnosed by emergency physicians. Given the high prevalence of sepsis in ICU settings (59), and in certain hospital institutions, an even higher prevalence of sepsis in emergency departments (45, 46), targeting ICU and emergency clinicians to improve documentation may lead to improvements in coding. Methods to improve the clinician documentation and the accuracy of sepsis diagnosis include the use of combination biomarkers (60), hospital performance improvement programs (58, 61), awareness campaigns (62), checklist-based bundles, and advanced electronic screening algorithms (63). Altogether, inadequate sepsis surveillance can impact the improvement and development of sepsis management programs, which would affect the outcomes of future patients (64).

Given the inaccuracy and significant variability of ICD-10 code sensitivity between different databases and centers, we recommend that internal validation of ICD-10 code characteristics (sensitivity, specificity) should be performed prior to using such codes for research, epidemiologic studies, and/or decision making. For example, data from chart reviews and/ or registry databases can be obtained and compared with a standardized sepsis definition (such as Sepsis-3 definitions [65]) to help assess the validity of ICD-10 codes the team wishes to use. Chart reviewers should be blinded to the corresponding chart's ICD-10 codes to decrease risk of bias. Although most studies in this systematic review employed retrospective methods to assess ICD-10 code validity, prospective collection and review of clinical and laboratory data prove an option as well. Researchers may consider accessing data from ongoing or pre-existing randomized control trials to identify patients with validated diagnoses/outcomes of interest. We also encourage investigators to delineate the type of sepsis evaluated (e.g., sepsis, severe sepsis, septic shock), and we highlight the need for assessment of ICD-10 coding among pediatric populations for which there is a paucity of data.

The strengths of this systematic review include the focus only on ICD-10 codes for identifying sepsis and a comprehensive search strategy that was not restricted to time, place, age, or inpatient/outpatient status. A careful evaluation of risk of bias using QUADAS and an overall certainty of evidence using GRADE were undertaken. Additionally, we conducted a meta-analysis to ascertain the pooled sensitivity and specificity of sepsis ICD-10 codes and secondary analyses to evaluate how study designs affect sepsis ICD-10 code accuracy.

There are several limitations to our review. First, we only included English language studies and did not perform a search on unpublished literature, although missing data are small and unlikely to change conclusions. Second, despite our broad search strategy, only two studies evaluated pediatric populations (37, 42) limiting generalizability to this population (41). Third, there was significant clinical and methodological heterogeneity since studies used different reference standard definitions, outcome definitions, study designs, clinical settings, coder quality (qualification, training, and years of experience), and ICD-10 codes studied. Fourth, given that the use of the terms sepsis, severe sepsis, and septicemia were not consistent across studies, we could not differentiate these terms in our analyses. Fifth, studies included in the meta-analysis were limited in methodology as they did not test ICD-10 accuracy against the 2017 Sepsis-3 definition (65) and did not include multicenter cohorts. Sixth, although one American abstract (Dunatchik et al [43]) was included in our systematic

review, no studies in the meta-analysis were conducted in the United States. Thus, our meta-analysis' sensitivity and specificity values do not fully reflect the potential bias from healthcare system payment incentives on ICD-10 sepsis codes (66).

The successor to ICD-10, 11th revision of ICD (ICD-11), is scheduled for release in January 2022 (67). The leap between coding versions will be significant, as ICD-11 will include more than triple the number of codes as ICD-10. Thus, future research should evaluate the accuracy of ICD-11 in identifying sepsis and whether there is a loss of coding quality during the transition (68).

CONCLUSIONS

Sepsis is underreported by ICD-10 coding, which may affect the perceived incidence of sepsis in clinical, claims, and administrative databases. Future research is needed to explore if greater consistency in ICD-10 code definitions and enhanced quality measures for ICD-10 coders can improve the diagnostic accuracy of sepsis from these large databases.

1 Department of Medicine, McMaster Centre for Transfusion Research, McMaster University, Hamilton, ON, Canada.

2 Department of Medicine, Michael G. DeGroote School of Medicine, McMaster University, Hamilton, ON, Canada.

3 Centre for Innovation, Canadian Blood Services, Ottawa, ON, Canada.

4 Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada.

5 Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada.

6 Canadian Blood Services, Ancaster, ON, Canada.

Bonnie Liu wrote the article. Bonnie Liu, Kayla J. Lucier, and Nancy M. Heddle coordinated the systematic review. Bonnie Liu and Anchit Garg developed the search strategy, and created the Grading of Recommendations, Assessment, Development, and Evaluation evidence profiles. Bonnie Liu, Anchit Garg, and Sophie Li screened and extracted the data, and evaluated studies for Quality Assessment of Diagnostic Accuracy Studies. Kayla J. Lucier and Dr. Ning acted as third-party adjudication during the screening and data extraction process. Melina Hadzi-Tosev and Yang Liu provided statistical data analysis. Melina Hadzi-Tosev, Kayla J. Lucier, Dr. Rochwerg, and Dr. Ning revised the article. Dr. Ning came up with the study idea. All authors approve of article submission and agree to be responsible for all aspects of work.

REFERENCES

1. Singer M, Deutschman CS, Seymour CW, et al: The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016; 315:801–810

2. Levy MM, Fink MP, Marshall JC, et al: 2001 SCCM/ESICM/ACCP/ATS/SIS international sepsis definitions conference. *Crit Care Med* 2003; 31:1250–1256

3. Rudd KE, Johnson SC, Agesa KM, et al: Global, regional, and national sepsis incidence and mortality, 1990-2017: Analysis for the global burden of disease study. *Lancet* 2020; 395:200–211

4. Organization WH, Others: Global Report on the Epidemiology and Burden of Sepsis: Current Evidence, Identifying Gaps and Future Directions, 2020. Available at https://apps.who.int/iris/bitstream/handle/10665/334216/9789240010789-eng.pdf. Access date: Sep 7, 2021

5. Rhodes A, Evans LE, Alhazzani W, et al: Surviving sepsis campaign: International guidelines for management of sepsis and septic shock: 2016. *Crit Care Med* 2017; 45:486–552

6. Monestime JP, Mayer RW, Blackwood A: Analyzing the ICD-10-CM transition and post-implementation stages: A public health institution case study. *Perspect Health Inf Manag* 2019; 16:1a

7. Butz J, Brick D, Rinehart-Thompson LA, et al: Differences in coder and physician perspectives on the transition to ICD-10-CM/PCS: A survey study [Internet]. *Health Policy Technol* 2016; 5:251–259

8. Cartwright DJ: ICD-9-CM to ICD-10-CM codes: What? Why? How? *Adv Wound Care* 2013; 2:588–592

9. Hennessy DA, Quan H, Faris PD, et al: Do coder characteristics influence validity of ICD-10 hospital discharge data? *BMC Health Serv Res* 2010; 10:99

10. Darbà J, Marsà A: Epidemiology, management and costs of sepsis in Spain (2008-2017): A retrospective multicentre study. *Curr Med Res Opin* 2020; 36:1089–1095

11. Suarez De La Rica A, Gilsanz F, Maseda E: Epidemiologic trends of sepsis in western countries. *Ann Transl Med* 2016; 4:325

12. Kadri SS, Rhee C, Strich JR, et al: Estimating ten-year trends in septic shock incidence and mortality in United States academic medical centers using clinical data. *Chest* 2017; 151:278–285

13. Fleischmann C, Thomas−Rueddel DO, Hartmann M, et al. Hospital Incidence and Mortality Rates of Sepsis: An Analysis of Hospital Episode (DRG) Statistics in Germany From 2007 to 2013 [Internet]. Deutsches Ärzteblatt international, 2016.

Available at: http://dx.doi.org/10.3238/arztebl.2016.0159. Accessed September 7, 2021

14. Duke GJ, Moran JL, Santamaria JD, et al: Sepsis in the new millennium - Are we improving? *J Crit Care* 2020; 56:273–280

15. Rhee C, Dantes R, Epstein L, et al: Incidence and trends of sepsis in us hospitals using clinical vs claims data, 2009-2014. *JAMA* 2017; 318:1241–1249

16. Fleischmann-Struzek C, Mikolajetz A, Schwarzkopf D, et al: Challenges in assessing the burden of sepsis and understanding the inequalities of sepsis outcomes between national health systems: Secular trends in sepsis and infection incidence and mortality in Germany. *Intensive Care Med* 2018; 44:1826–1835

17. Quan H, Li B, Saunders LD, et al: Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res* 2008; 43:1424–1441

18. Chute CG, Huff SM, Ferguson JA, et al: There are important reasons for delaying implementation of the new ICD-10 coding system. *Health Aff (Millwood)* 2012; 31:836–842

19. Gedeborg R, Furebring M, Michaëlsson K: Diagnosis-dependent misclassification of infections using administrative data variably affected incidence and mortality estimates in ICU patients. *J Clin Epidemiol* 2007; 60:155–162

20. Jolley RJ, Sawka KJ, Yergens DW, et al: Validity of administrative data in recording sepsis: A systematic review. *Crit Care* 2015; 19:139

21. Barber C, Lacaille D, Fortin PR: Systematic review of validation studies of the use of administrative data to identify serious infections. *Arthritis Care Res* 2013; 65:1343–1357

22. Jetté N, Quan H, Hemmelgarn B, et al: The development, evolution, and modifications of ICD-10: Challenges to the international comparability of morbidity data. *Med Care* 2010; 48:1105–1110

23. Nichols JC, Osmani FA, Sayeed Y: Importance of proper utilization of international classification of diseases 10th revision and clinical documentation in modern payment models. *J Arthroplasty* 2016; 31:945–946

24. McInnes MDF, Moher D, Thombs BD, et al; and the PRISMA-DTA Group: Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *JAMA* 2018; 319:388–396

25. Whiting P, Rutjes AWS, Reitsma JB, et al: The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3:25

26. Ryan R, Hill S: How to GRADE the Quality of the Evidence [Internet]. Cochrane Consumers and Communication Group. UK, 2016. Available at: https://cccrg.cochrane.org/author-resources. Accessed September 7, 2021

27. DerSimonian R, Laird N: Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7:177–188

28. Jackson D: Confidence intervals for the between-study variance in random effects meta-analysis using generalised Cochran heterogeneity statistics. *Res Synth Methods* 2013; 4:220–229

29. R: A Language and Environment for Statistical Computing [Internet]. R Foundation for Statistical Computing. Vienna, Austria. Available at https://paperpile.com/app/p/baec1401-c8ac-0b66-8c7f-6d97d798d757. Accessed January 25, 2022

30. Higgins JPT Green S: Cochrane Handbook for Systematic Reviews of Interventions [Internet]. The Cochrane Collaboration.UK, 2011. Available at: www.handbook.cochrane.org. Accessed September 7, 2021

31. Søgaard KK, Thomsen RW, Schønheyder HC, et al: Positive predictive values of the international classification of diseases, 10th revision diagnoses of Gram-negative septicemia/sepsis and urosepsis for presence of Gram-negative bacteremia. *Clin Epidemiol* 2015; 7:195–199

32. Lauridsen MD, Gammelager H, Schmidt M, et al: Positive predictive value of international classification of diseases, 10th revision, diagnosis codes for cardiogenic, hypovolemic, and septic shock in the Danish national patient registry. *BMC Med Res Methodol* 2015; 15:23

33. Madsen KM, Schønheyder HC, Kristensen B, et al: Can hospital discharge diagnosis be used for surveillance of bacteremia? A data quality study of a Danish hospital discharge registry. *Infect Control Hosp Epidemiol* 1998; 19:175–180

34. Holland-Bill L, Xu H, Sørensen HT, et al: Positive predictive value of primary inpatient discharge diagnoses of infection among cancer patients in the Danish national registry of patients [Internet]. *Ann Epidemiol* 2014; 24:593–7, 597.e1

35. Ibrahim I, Jacobs IG, Webb SAR, et al: Accuracy of international classification of diseases, 10th revision codes for identifying severe sepsis in patients admitted from the emergency department. *Crit Care Resusc* 2012; 14:112–118

36. Das A, Kennedy K, Spyropoulos G, et al: Administrative data has poor accuracy for surveillance of Staphylococcus aureus bacteraemia. *Infect Dis Health* 2016; 21:162–168

37. Reilly JR, Shulman MA, Gilbert AM, et al: Towards a national perioperative clinical quality registry: The diagnostic accuracy of administrative data in identifying major postoperative complications. *Anaesth Intensive Care* 2020; 48:203–212

38. Fleischmann-Struzek C, Thomas-Rüddel DO, Schettler A, et al: Comparing the validity of different ICD coding abstraction strategies for sepsis case identification in German claims data. *PLoS One* 2018; 13:e0198847

39. Quan H, Eastwood C, Cunningham CT, et al: Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). *BMJ Open* 2013; 3:e003716

40. Jolley RJ, Quan H, Jetté N, et al: Validation and optimisation of an ICD-10-coded case definition for sepsis using administrative health data. *BMJ Open* 2015; 5:e009487

41. Parthasarathy M, Reid V, Pyne L, et al: Are we recording postoperative complications correctly? Comparison of NHS hospital episode statistics with the American College of Surgeons National Surgical Quality Improvement Program. *BMJ Qual Saf* 2015; 24:594–602

42. Agyeman PKA, Endrich O, Triep K, et al: Accuracy of ICD-10 coding for sepsis and organ dysfunctions in children with blood culture-proven sepsis in Switzerland, results form the prospective Swiss pediatric sepsis study. Conference: Annual Meeting Swiss Society of Paediatrics Bellinzona [Internet]. Bellinzona, Switzerland, 2019. June 6,7 2019. Available at https://www.researchgate.net/publication/333772392_Accuracy_of_ICD-10_coding_for_sepsis_and_organ_dysfunctions_in_children_with_blood_culture-proven_sepsis_in_Switzerland_results_from_the_prospective_Swiss_Pediatric_Sepsis_Study. Accessed September 7, 2021

43. Dunatchik A, Semler MW, Rice TW, et al: Accuracy of the centers for medicare and medicaid services ICD-10-CM codes in identifying sepsis among critically ill adults. *Am J Respir Crit Care Med* 2017: A5016. Available at: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2017.195.1_MeetingAbstracts.A5016. Accessed October 21, 2022

44. Henriksen DP, Laursen CB, Jensen TG, et al: Incidence rate of community-acquired sepsis among hospitalized acute medical patients-a population-based survey. *Crit Care Med* 2015; 43:13–21

45. Yu CW, Chang SS, Lai CC, et al: Epidemiology of emergency department sepsis: A national cohort study between 2001 and 2012. *Shock* 2019; 51:619–624

46. Oberlin M, Balen F, Bertrand L, et al: Sepsis prevalence among patients with suspected infection in emergency department: A multicenter prospective cohort study. *Eur J Emerg Med* 2020; 27:373–378

47. Wang HE, Addis DR, Donnelly JP, et al: Discharge diagnoses versus medical record review in the identification of community-acquired sepsis. *Crit Care* 2015; 19:42

48. Whittaker SA, Mikkelsen ME, Gaieski DF, et al: Severe sepsis cohorts derived from claims-based strategies appear to be biased toward a more severely ill patient population. *Crit Care Med* 2013; 41:945–953

49. Surjan G: Questions on validity of international classification of diseases-coded diagnoses. *Int J Med Inform* 1999; 54:77–95

50. Genga KR, Russell JA: Update of sepsis in the intensive care unit. *J Innate Immun* 2017; 9:441–455

51. Drösler SE, Romano PS, Sundararajan V, et al: How many diagnosis fields are needed to capture safety events in administrative data? Findings and recommendations from the WHO ICD-11 topic advisory group on quality and safety. *Int J Qual Health Care* 2014; 26:16–25

52. O'Malley KJ, Cook KF, Price MD, et al: Measuring diagnoses: ICD code accuracy. *Health Serv Res* 2005; 40:1620–1639

53. White J, Jawad M: Clinical coding audit: No coding - no income - no hospital. *Cureus* 2020; 12:e10664

54. Possel ML: "International Classification of Diseases, Tenth Revision, Clinical Modification and Procedure Coding System and Clinical Documentation Improvement. Applied Research Projects [Internet] . University of Tennessee Health Science Center, 2014. Available at http://dx.doi.org/10.21007/chp.hiim.0038. Accessed September 7, 2021

55. Shepheard J: Clinical coding and the quality and integrity of health data. *Health Inf Manag* 2020; 49:3–4

56. Bedoya AD, Futoma J, Clement ME, et al: Machine learning for early detection of sepsis: An internal and temporal validation study. *JAMIA Open* 2020; 3:252–260

57. Dash S, Shakyawar SK, Sharma M, Kaushik S. Big data in healthcare: Management, analysis and future prospects. *J Big Data [Internet]* 2019; 6:54

58. Rezende E, Silva JM Jr, Isola AM, et al: Epidemiology of severe sepsis in the emergency department and difficulties in the initial assistance. *Clinics (Sao Paulo)* 2008; 63:457–464

59. Sakr Y, Jaschinski U, Wittebole X, et al: Sepsis in intensive care unit patients: Worldwide data from the intensive care over nations audit. *Open Forum Infect Dis* 2018; 5:ofy313

60. Vincent JL: The clinical challenge of sepsis identification and monitoring. *PLoS Med* 2016; 13:e1002022

61. Kim HI, Park S: Sepsis: Early recognition and optimized treatment. *Tuberc Respir Dis (Seoul)* 2019; 82:6–14

62. Fiest KM, Krewulak KD, Brundin-Mather R, et al: Patient, public, and healthcare professionals' sepsis awareness, knowledge, and information seeking behaviors: A scoping review. *Crit Care Med* 2022; 50:1187–1197

63. Graber ML, Patel M, Claypool S: Sepsis as a model for improving diagnosis. *Diagnosis (Berl)* 2018; 5:3–10

64. Amland RC, Sutariya BB: Quick sequential [sepsis-related] organ failure assessment (qSOFA) and St. John sepsis surveillance agent to detect patients at risk of sepsis: An observational cohort study. *Am J Med Qual* 2018; 33:50–57

65. Seckel MA: Sepsis-3. *Nurs Crit Care* 2017; 12:37–43

66. Kesselheim AS, Brennan TA: Overbilling vs. downcoding--the battle between physicians and insurers. *N Engl J Med* 2005; 352:855–857

67. World Health Organization. ICD-11: Classifying Disease to Map the Way We Live and Die [Internet]. World Health Organization. Available at https://web.archive.org/web/20180620014204/https://www.who.int/health-topics/international-classification-of-diseases. Accessed September 7, 2021

68. ICD 11 Timeline [Internet], World Health Organization. Available at: https://web.archive.org/web/20220121042954/https://www.who.int/classifications/icd/revision/timeline/en/. Accessed September 7, 2021