Supplementary data

"Accuracy of administrative data for surveillance of healthcare-associated infections: a systematic review"

Authors:

Maaike S.M. van Mourik, Pleun Joppe van Duijn, Karel G.M. Moons, Marc J.M. Bonten, Grace M. Lee

Contents:

S1. Search Strategy

Databases: Medline/Pubmed, EMBASE, CINAHL, Cochrane.
All searches in Titles + Abstract
Limits: Published between after 1995, Languages: English, Dutch, French, German
Search dates: Initial search march 8th 2012, search closure March 1st 2013.

Outcome: Healthcare associated infection	Search terms :
	Infection, infections, hai, infectious, sepsis, meningitis, notifiable, SSI, VAP, pneumonia, CAUTI, CLABSI, CABSI, BSI

AND

Determinant: administrative data	Search terms :
	ICD, international Classification of Diseases, administrative, discharge diagnos*, registry, registries, electronic data, claim data, claims data, reimbursement, health plan data, healthplan, medicare, diagnostic coding, discharge coding, discharge code(s), diagnostic coding, diagnostic code(s), diagnosis code(s), diagnosis coding, procedure code(s), procedure coding

S2. Data collection, quality assessment items and assumptions

General characteristics

Item	Options	Considerations & assumptions
Author, year of publication		
HAI studied	SSI/BSI/sepsis/ CLABSI/VAP /UTI/CAUTI/Other	More than 1 may apply Specify details
Systematic post-discharge surveillance?	Yes/No	Only code as yes if explicit aim of the study.
Location of study	Country	
Number of participating centers		
Start and stop of patient inclusion		
Validation of previously developed algorithm	Yes/No	E.g. previous study, PHC4, PSI, HAC
Validation sample within the study	Yes/No	
Purpose of administrative data	Billing/ benchmarking /demographic/ unclear	If U.S.: code as billing
Setting: Medicare, VA or HMO only?	Yes/No (specify)	
Healthcare setting	Primary care, Inpatient, Outpatient, ICU	More than 1 possible
Academic hospital	Yes/No/Mixed (if multicenter)	
Public reporting	Yes/Potentially/No	Was the measure developed/tested as a means of public reporting or external quality benchmarking (as opposed to an in-hospital screening algorithm)

Assessment of risk of bias (adapted from QUADAS-2)

	PATIENT SELECTION		
1	Method of patient selection	Describe in-/exclusion criteria	
2	Consecutive or random sample of patients enrolled	Yes/no	Random sampling scored as yes
3	Case-control design avoided	Yes/No	
4	Inappropriate exclusions avoided?	Yes/No	Is the sample enrolled representative of the domain (e.g. no exclusion of high-risk patients?)
5	Risk of bias patient selection	Low/Unclear/High	If#2, #3 or #4 = no, consider risk of bias
6	Applicability patient selection	Low/Unclear/High	
	INDEX TEST		
1	Describe index test	Coding system used? Codes assigned by? Procedure codes to detect HAI? PSI algorithm List codes used, duration of follow-up	ICD-9 or ICD-10 Coders, physicians, other, unclear (US: professional coders assumed) No if only used to identify patients at risk Version number Specify use of pre-defined methods (PHC4, PSI, CMS…).
2	Were other tests assessed	Yes/No, specify	
3	Was the administrative data intended as the sole method of surveillance	Yes/no	E.g. were results of administrative data intended to be combined with microbiology results?
4	Was interpretation done without knowledge of the reference standard?	Yes/no	Were codes assigned without knowledge of reference standard?
5	Pre-specified threshold	Yes/no	Was code selection determined in advance? If unspecified and only a very specific code is used, also code as yes (e.g. 998.5 for SSI)
6	Risk of bias index test	Low/Unclear/High	If #4 or #5 = No, consider risk of bias.
7	Applicability index test	Low/Unclear/High	If #3 = No, score as High
	REFERENCE STANDARD		
1	Describe reference standard	Method: Definitions used: Applied by:	Describe NHSN/NNIS, (VA)SQIP, Clinical, Other IP, trained nurses, physicians, other abstractor

2	Is the reference standard likely to correctly classify the patient	Yes/No	
3	Was it interpreted without knowledge of the index test?	Yes/No	If only patients flagged by code are received reference standard and/or coding status was unblinded score as No
4	Risk of bias	Low/Unclear/High	If #3 = No, consider risk of bias
5	Applicability	Low/Unclear/High	
FLOW AND TIMING			
1	Describe patients who did not receive 1 of both tests or are not in 2x2 table		Draw flowchart
2	Did all patients receive the RS?	Yes/No	If only assessing patients with positive reference test, score as No
3	Did all patients receive the same RS?	Yes/No	If all the patients receiving RS do not receive the *same* RS score as No.
4	Were all patients included in the analysis?	Yes/No	
5	Could the patient flow have introduced bias and why?	Low/Unclear/High	If #2 or #3 = Yes, consider risk of bias. If a large or important portion of patients are excluded (e.g. due to missing data), consider risk of bias.
6	How were missing data handled?	Description	

Data extraction:

	HAI present	HAI absent	Total
Codes +	TP	FP	
Codes -	FN	TN	
Total			

If only outcome measures are reported:

Sensitivity		PPV	
Specificity		NPV	
LR-		LR+	
Kappa		Degree of certainty	High – med – low

General remarks:
- If multiple index tests and/or reference standards and/or patient flow schemes are used in the study, all are assessed separately for their risk of bias (multiple comparisons).
- Data were extracted for each comparison presented, and also separately if
    o Multiple types of HAI
    o Multiple comparisons for each HAI
    o If multiple specifications of administrative data

S3. Risk of bias individual studies, stratified in case of multiple comparisons

Abbreviations & Legend

HAI types: (CA)UTI – (catheter-associated) urinary tract infection, (CLA)BSI – central-line associated bloodstream infection, Pneu – pneumonia, SSI – surgical site infection, VAP – ventilator-associated pneumonia.

Country: AUS – Australia, B   E – Belgium, CAN – Canada, DK – Denmark, ESP – Spain, FI – Finland, FR – France, IT-Italy, JP – Japan, NL – Netherlands, USA – United States of America,

Definition: CDC-NHSN or CDC-NNIS – definitions from the Centers for Disease Control Healthcare Safety Network or its predecessor, (VA/N)SQIP – definitions & methods from the National (or Veteran's Affairs) Surgical Quality Improvement Project.

Intend appl: Intended application of administrative data within HAI surveillance.

Ext – for external quality assessment, e.g. public reporting or pay-for-performance.

Int (S) – to support within hospital surveillance as sole method of finding possible HAI cases.

Int (C) – to support within hospital surveillance, combined with other indicators of HAI.

If applicable, specific metrics are indicated: HAC – Healthcare-associated condition as defined by the Centers for Medicare and Medicaid Services, PHC4 – code selection specified by the Pennsylvania Healthcare Cost Containment Council, PSI – Patient Safety Indicator.

N : design number

Risk of bias (Rob) & applicability domains: Patient selection (Pat Sel), Index test, Reference standard (Ref) and Flow. If a study assesses only the positive predictive value (partial verification, fully dependent on the index test – e.g. administrative data), and the risk of bias of the on the flow domain is low for the PPV estimate, these studies have been marked as "PPV" in the risk of bias on flow column. The overall risk of bias of the PPV estimate is marked in RoB PPV column.

Notes:

The following studies used the ICD-10 coding system: Curtis 2004, Daneman 2011, Gerbier 2011, Kanerva 2009, Lee 2011, Leth 2006, Leth 2010. Heisler 2009 used a different coding system.

In the following studies a present-on-admission indicator was explicitly included in the administrative data algorithm:

Cima 2011, Haley 2012, Koch 2012, Meddings 2010, Moehring 2013, Murff 2011, Tehrani 2013, Zrelak 2011

Author & year	HAI studied	Country	N Centers	Study period	definition	Intend appl	N	Risk of bias			Applicability				RoB PPV
								Pat sel	Index test	Ref	Flow	Pat sel	Index test	Ref	
Apte, 2011	SSI,	USA	1	2007	Unclear	Int (C)	2	Low	Low	High	Low	Low	Low	High	High
Apte, 2011	SSI,	USA	1		CDC NHSN	Int (C)		Low	Low	Low	Uncl	Low	Low	Low	Uncl
Best, 2002	SSI, Sepsis, Pneu, UTI,	USA	123	1994 - 1995	(VA/N)SQIP	Ext	1	Uncl	Low	Low	Low	Low	Low	Low	Uncl
Bolon, 2009	SSI,	USA	8	2002 - 2005	CDC NHSN	Int (C)	1	Low	Low	Uncl	Low	Low	High	Low	Low
Braun, 2006	BSI,	USA	28	1999	Clinical	Ext*	1	Uncl	Low	High	High	Low	Low	High	High
Cadwallader, 2001	SSI,	AUS	1	1998 - 1999	CDC NNIS	Int (S)	2	Low	Low	Low	Low	Low	Low	Low	Low
Cadwallader, 2001	SSI,	AUS	1		CDC NNIS	Int (S)		Low	Low	High	High	Low	Low	Low	High
Calderwood, 2012	SSI,	USA	4	2007	CDC NHSN	Int (S)	1	Uncl	Uncl	High	High	Low	Low	Low	High
Calderwood, 2013	SSI,	USA	3296	2005 - 2007	CDC NHSN	Ext	2	Low	Low	High	PPV	Low	Low	Low	High
Calderwood, 2013	SSI,	USA	3296		CDC NHSN	Ext		Low	High	High	PPV	Low	Low	Low	High
Campbell, 2011	SSI, UTI,	USA	1	2008	Other	Int (S)	1	Uncl	Uncl	Low	Low	Low	Low	High	Uncl
Cevasco, 2011a	CLABSI,	USA	28	2002 - 2007	Other	Ext PSI 3.1	1	Low	Low	High	PPV	Low	Low	Low	High
Cevasco, 2011b	Sepsis,	USA	75	2003 - 2007	Other	Ext PSI 3.1	2	Low	Low	High	PPV	Low	Low	Low	High
Cevasco, 2011b	Sepsis,	USA	75		Unclear	Ext PSI 3.1		Low	Low	High	PPV	Low	Low	Low	High
Cima, 2011	CLABSI, Sepsis,	USA	1	2006 - 2009	(VA/N)SQIP	Ext PSI 3.1	1	Low	Low	Low	Low	Low	Low	Low	Low
Curtis, 2004	SSI,	AUS	1	2001 - 2002	Other	Int (S)	2	Low	Low	Low	Low	Low	Low	Low	Low
Curtis, 2004	SSI,	AUS	1		Other	Int (S)		Low	Low	Uncl	High	Low	Low	Low	High
Daneman, 2011	SSI,	CAN	1	2008 - 2009	CDC NHSN	Int (S)	1	Uncl	Low	Low	Low	Low	Low	Low	Uncl
Drees, 2010	VAP,	USA	1	2007 - 2008	CDC NHSN	Int (S)	1	Low	Low	Low	Uncl	Low	Low	Low	Low
Gerbier, 2011	SSI, BSI, CLABSI, UTI, Pneu,	FR	1	2000 - 2007	Other	Int (S)	1	Low	Low	Low	Uncl	Low	Low	Low	Low
Haley, 2012	SSI,	USA	176	2008 - 2010	CDC NHSN	Ext	2	Low	Uncl	Low	Low	Low	Low	Low	Low
Haley, 2012	SSI,	USA	176		CDC NHSN	Ext		Low	Uncl	High	High	Low	Low	Low	High
Hebden, 2000	SSI,	USA	1	1997	CDC NNIS	Int (S)	1	Low	Low	Low	Low	Low	Low	Low	Low
Heisler, 2009	UTI, CAUTI,	USA	1	2004 - 2005	Clinical	Int (S)	1	Low	Low	High	Uncl	Low	Low	Uncl	High
Hollenbeak, 2011	SSI,	USA	1	2007 - 2008	CDC NHSN	Int (S)	1	Low	Low	Low	Low	Low	Low	Low	RoB
Hougland, 2008	BSI, Pneu	USA	77	2001 - 2003	Unclear	Ext	1	Low	Low	Low	Uncl	Low	Low	Uncl	Low

Author & year	HAI studied	Country	N Centers	Study period	definition	Intend appl	N	Risk of bias				Applicability			RoB PPV
								Pat sel	Index test	Ref	Flow	Pat sel	Index test	Ref	
Huang, 2011	SSI,	USA	671	2005	CDC NHSN	Ext	3	Low	High	High	High	Low	Low	Low	High
Huang, 2011	SSI,	USA	671		Unclear	Ext		Low	Low	High	Uncl	Low	Low	High	High
Huang, 2011	SSI,	USA	671		CDC NHSN	Ext		Low	Low	High	High	Low	Low	Low	High
Inacio, 2011	SSI,	USA	?	2006 - 2008	CDC NHSN	Int (S)	1	Low	Low	Low	Low	Low	Low	Low	Low
Julian, 2006	SSI, VAP, UTI, CAUTI,	USA	1	2004	CDC NHSN	Ext PHC4	1	Low	Low	High	High	Low	Low	Low	High
Kanerva, 2009	SSI, BSI, UTI, Pneu,	FI	20	2005	Other	Int (S)	1	Low	Uncl	Low	Low	Low	Low	Low	Uncl
Koch, 2012	Sepsis,	USA	1	2009 - 2010	(VA/N)SQIP	Ext PSI 4.2	2	Low	Low	Low	Low	Low	Low	Low	Low
Koch, 2012	Sepsis,	USA	1		Other	Ext PSI 4.2		Low	Low	Low	Low	Low	Low	Low	Low
Landers, 2010	UTI,	USA	1	2007	Other	Int (S)	1	Low	Low	Low	Low	Low	Low	High	Low
Lawson, 2012	SSI, Sepsis, Pneu, UTI,	USA	214	2005 - 2008	(VA/N)SQIP	Ext	1	Low	Uncl	Low	Low	Low	Low	Low	Uncl
Lee, 2011	SSI, BSI, Pneu, UTI,	JP	4	2005 - 2009	CDC NHSN	Int (C) PHC4	1	Low	Low	Low	Low	High	High	Low	Low
Leth, 2006	SSI,	DK	1	1999 - 2002	CDC NHSN	Int (C)	2	Low	Uncl	Low	Low	Low	High	Low	Low
Leth, 2006	SSI,	DK	1	1999 - 2002	CDC NHSN	Int (C)		Uncl	Low	Uncl	Low	Low	Low	Low	Uncl
Leth, 2010	SSI	DK	3	2007 - 2008	CDC NHSN	Int (C)	1	Low	Low	Low	High	Low	High	Low	High
Meddings, 2010	CAUTI,	USA	1	2006 - 2007	Other	Ext HAC	1	Low	Low	High	High	Low	Low	High	High
Miner, 2004	SSI,	USA	7	1996 - 1999	CDC NNIS	Int (C)	1	Low	Low	High	High	Low	High	Low	High
Moehring, 2013	CLABSI,	USA	3	2007 - 2009	CDC NHSN	Ext HAC	1	Low	Low	Low	High	Low	Low	Low	Low
Moro, 2004	SSI,	IT	31	2001	CDC NNIS	Int (S)	1	Low	Uncl	Low	Low	Low	Low	Low	Uncl
Murff, 2011	Sepsis, Pneu	USA	6	1999 - 2006	(VA/N)SQIP	Ext PSI 3.1	1	Low	Low	Low	Low	Low	Low	Low	Low
Ollendorf, 2002	Sepsis,	USA	10	Uncl	Clinical	Int (S)	1	Uncl	Uncl	Low	Low	Uncl	Low	High	Uncl
Olsen, 2010	SSI,	USA	1	1998 - 2002	CDC NHSN	Int (S)	1	Uncl	Low	High	High	Low	Low	Low	High
Platt, 2002	SSI,	USA	4	1996 - 1999	CDC NNIS	Int (C)	1	Uncl	Low	High	High	Low	High	Low	High
Pokorny, 2006	CLABSI, VAP, CAUTI,	ESP	1	1999 - 2002	CDC NHSN	Int (C)	1	Low	Uncl	Low	Low	Low	High	Uncl	Uncl
Romano, 2009	Sepsis,	USA	110	2000 - 2001	(VA/N)SQIP	Ext PSI 2.1	2	Low	Low	Low	Low	Low	Low	Low	Low
Romano, 2009	Sepsis,	USA	110	2000 - 2001	(VA/N)SQIP	Ext PSI 2.1		Low	High	Low	Low	Low	Low	Low	High
Sands, 2003	SSI,	USA	5	1995 - 1997	CDC NNIS	Int (C)	1	Uncl	Low	High	High	Low	High	Low	High

| Author & year	HAI studied	Country	N Centers	Study period	definition	Intend appl	N	Risk of bias				Applicability			RoB PPV
								Pat sel	Index test	Ref	Flow	Pat sel	Index test	Ref	
Scanlon, 2008	CLABSI, Sepsis,	USA	28	2003 - 2005	Other	Ext PDI	1	Low	Low	High	PPV	Low	Low	High	High
Sherman, 2006	SSI, CLABSI, VAP, CAUTI,	USA	1	2004	CDC NHSN	Ext PHC4	1	Low	Low	High	High	Low	Low	Low	High
Song, 2008	SSI,	USA	1	2005	CDC NNIS	Int (C)	1	Low	Uncl	Low	Low	Low	High	Low	Uncl
Spolaore, 2005	SSI,	IT	3	2001	CDC NHSN	Int (C)	1	Low	Low	High	PPV	Low	High	Low	High
Stamm, 2012	CLABSI, VAP, CAUTI,	USA	1	2009	CDC NHSN	Int (S)	1	Low	Uncl	Uncl	High	Low	Low	Low	High
Stevenson, 2008	SSI, CLABSI, VAP,	USA	1	2005	CDC NHSN	Ext PHC4	2	Low	Low	Low	Low	Low	Low	Low	Low
Stevenson, 2008	SSI, CLABSI, VAP,	USA	1	2005	CDC NHSN	Ext PHC4		Low	Low	Uncl	High	Low	Low	Low	High
Stone, 2007	CLABSI,	USA	24	2002	CDC NHSN	Ext PSI 2.1	1	Low	Low	Low	Low	Low	Low	Low	Low
Tehrani, 2013	CLABSI,	USA	6	2009 - 2011	CDC NHSN	Ext HAC	2	Low	Low	Low	Low	Low	Low	Low	Low
Tehrani, 2013	CLABSI,	USA	6	2009 - 2011	CDC NHSN	Ext HAC		Low	Low	Uncl	PPV	Low	Low	Low	Low
Tinelli, 2011	SSI, UTI,	USA	28	2005 - 2006	CDC NHSN	Int (S)	1	Low	Uncl	Low	Low	Low	Low	Low	Uncl
van Mourik, 2013	Drain-related meningitis	NL	1	2004 - 2010	CDC NHSN	Int (S)	1	Uncl	Low	Low	Low	Low	Low	Low	Uncl
Verelst, 2010	SSI, Sepsis, VAP,	BE	8	2005	Clinical	Ext PSI 3.1	1	High	Low	Low	Uncl	Low	Low	Low	High
Yokoe, 2001	Postpartum	USA	1	1993 - 1995	CDC NNIS	Int (C)	1	Low	Low	High	High	Low	High	Low	High
Yokoe, 2004	SSI,	USA	13	1998 - 2001	CDC NNIS	Int (C)	2	Low	Low	High	High	Low	High	Low	High
Yokoe, 2004	SSI,	USA	13	1998 - 2001	CDC NNIS	Int (C)		Low	Low	High	Uncl	Low	High	Low	High
Yokoe, 2012	SSI,	USA	5	2003 - 2005	CDC NHSN	Int (C)	1	Low	Low	Uncl	Low	Low	High	Low	Low
Zhan, 2009	CAUTI,	USA	uncl	2005 - 2006	Other	Ext	1	Uncl	Uncl	Low	Low	Low	Low	Uncl	Uncl
Zrelak, 2011	CLABSI,	USA	23	2005	CDC NHSN	Ext PSI 3.1	1	Low	Low	High	PPV	Low	Low	Low	High

S4. Summary risk of bias, by HAI type.

Risk of bias was assessed using the *Quality Assessment for Diagnostic Accuracy Studies (QUADAS-2)* methods. Some studies contain multiple comparisons; in this case the lowest risk of bias per study is included. Shading denotes studies where extraction of complete two-by-two tables was not possible, including studies only assessing positive predictive values.
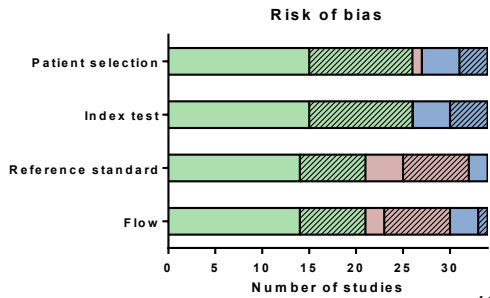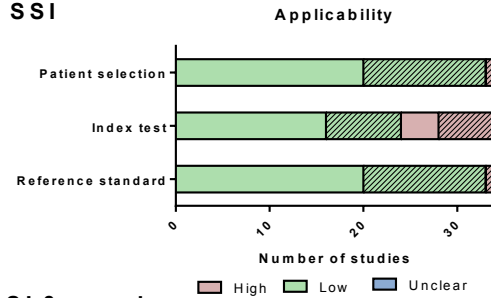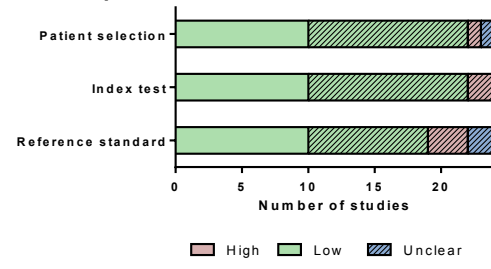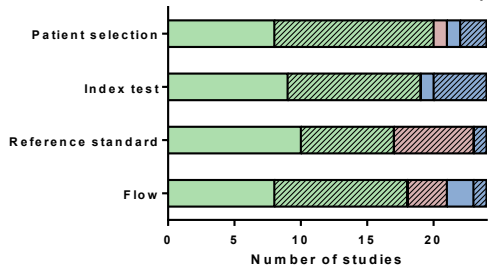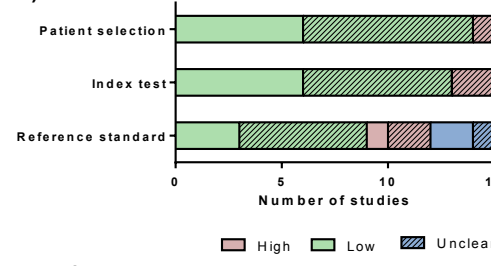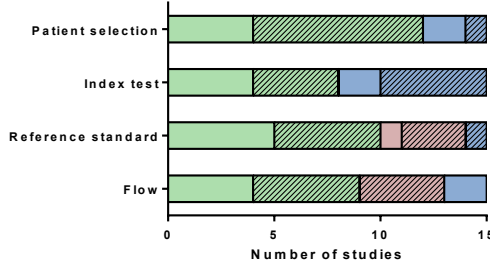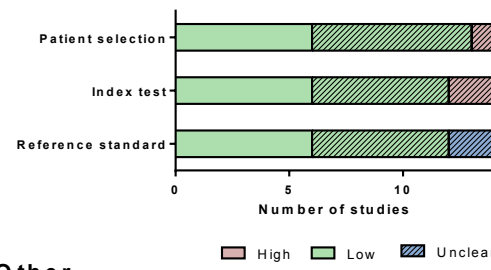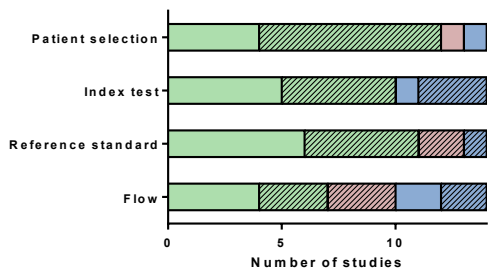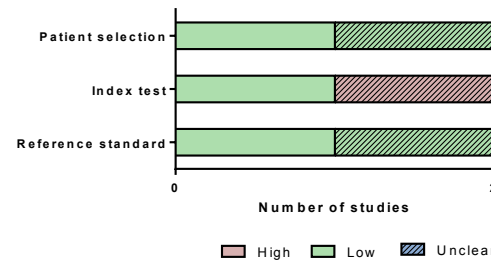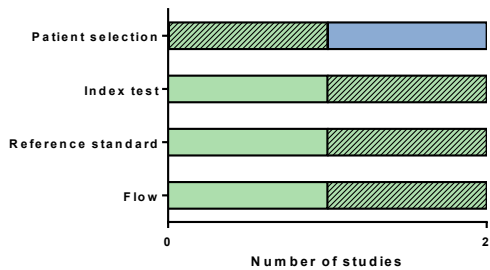
Figure S5. Forest plots for specificity and negative predictive value, stratified by HAI type and relevant study characteristics.

Studies are grouped by the intended application of administrative data:

Int (S) – used in isolation to support within-hospital surveillance efforts,

Int (C) – used to support within-hospital surveillance, combined with other indicators of infection,

Ext – for external quality assessment, including public reporting and pay-for-performance.
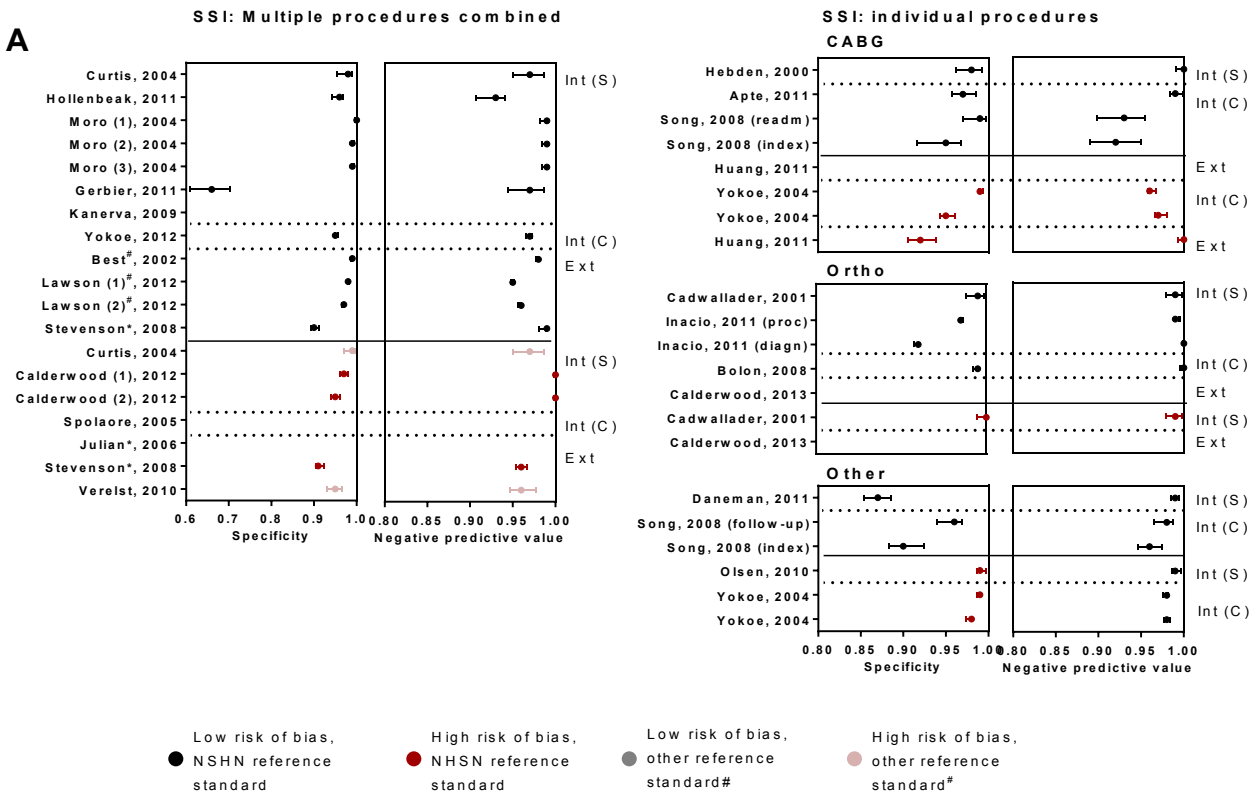
In studies including multiple specifications of the administrative data algorithm, these are numbered sequentially. 95% confidence intervals are derived using the exact binomial method.

Abbreviations: BSI – bloodstream infection, CABG – coronary artery bypass graft, DRM – drain-related meningitis, Ortho – orthopedic Procedure, PSI – Patient safety indicator, Sep – Sepsis, SSI – surgical site infection, UTI – urinary tract infection.

#: reference standard from Surgical Quality ImProvement Project (NSQIP or VASQIP). *: Code selection based on specification from Pennsylvania Health Cost Containment Council. **: HAC specification.

A. Surgical site infection, B. (Catheter-associated) bloodstream infection, C. (Catheter-associated) urinary tract infection, D. (Ventilator-associated) pneumonia. E. Other HAI or studies Extesenting only data aggregated for multiple types of infection.

B

Central line associated BSI

Cima[#], 2011
Stone, 2007
Stevenson*, 2008
Moehring**, 2013
Tehrani**, 2013
Stamm, 2012
Cevasco, 2011a
Scanlon, 2008
Zrelak, 2011
Stevenson*, 2008
Moehring**, 2013

Ext (PSI 7)
Ext
Int (S)
Ext (PSI 7)
Ext

Specificity | Negative predictive value
0.6 0.7 0.8 0.9 1.0 | 0.80 0.85 0.90 0.95 1.00

● Low risk of bias, NHSN reference standard
● Low risk of bias, other reference standard[#]
● High risk of bias, NHSN reference standard
● High risk of bias, other reference standard[#]

Bloodstream infection (BSI)

Gerbier, 2011
Kanerva, 2009
Ollendorf, 2002
Cima[#], 2011
Koch[#], 2012
Murff[#], 2011
Romano[#], 2009
Best[#], 2002
Hougland, 2008
Lawson (1)[#], 2012
Lawson (2)[#], 2012
Cevasco, 2011b (VA)
Cevasco, 2011b (com)
Romano[#], 2009
Scanlon, 2008
Verelst, 2010

Int (S)
Ext (PSI 13)
Ext
Ext (PSI 13)

Specificity | Negative predictive value
0.80 0.85 0.90 0.95 1.00 | 0.80 0.85 0.90 0.95 1.00

C

Catheter-associated UTI

Zhan (1), 2009
Zhan (2), 2009
Zhan (3), 2009
Stamm, 2012
Meddings**, 2010

Ext
Int (S)
Ext

Specificity | Negative predictive value
0.80 0.85 0.90 0.95 1.00 | 0.80 0.85 0.90 0.95 1.00

● Low risk of bias, NSHN reference standard
● High risk of bias, NHSN reference standard
● Low risk of bias, other reference standard[#]
● High risk of bias, other reference standard[#]

Urinary tract infection

Gerbier, 2011
Kanerva, 2009
Landers, 2010
Best[#], 2002
Lawson (1)[#], 2012
Lawson (2)[#], 2012
Heisler, 2009
Julian, 2006

Int (S)
Ext
Int (S)
Ext

Specificity | Negative predictive value
0.80 0.85 0.90 0.95 1.00 | 0.80 0.85 0.90 0.95 1.00

D

Ventilator-associated pneumonia

Drees, 2010
Stevenson*, 2008
Stamm, 2012
Julian*, 2006
Stevenson*, 2008
Verelst, 2010

Int (S)
Ext
Int (S)
Ext

Specificity | Negative predictive value
0.70 0.80 0.90 1.00 | 0.80 0.85 0.90 0.95 1.00

● Low risk of bias, NSHN reference standard
● High risk of bias, NHSN reference standard
● Low risk of bias, other reference standard[#]
● High risk of bias, other reference standard[#]

Pneumonia

Gerbier, 2011
Kanerva, 2009
Best[#], 2002
Lawson (1)[#], 2012
Lawson (2)[#], 2012
Murff[#], 2011
Hougland, 2008

Int (S)
Ext

Specificity | Negative predictive value
0.80 0.85 0.90 0.95 1.00 | 0.80 0.85 0.90 0.95 1.00

E

Other HAI and aggregate estimates

Gerbier, 2011
Kanerva, 2009
Tinelli, 2011
van Mourik, 2013
Pokorny, 2006
Lee*, 2011
Sherman*, 2006

Int (S)
Int (C)
Ext

Specificity | Negative predictive value
0.80 0.85 0.90 0.95 1.00 | 0.80 0.85 0.90 0.95 1.00

● Low risk of bias, NHSN reference standard
● Low risk of bias, other reference standard[#]
● High risk of bias, NHSN reference standard
● High risk of bias, other reference standard[#]

11

Table S6: Administrative data algorithm, by HAI type

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

SSI – CABG

Study	Codes used (Inpatient only, primary & secondary codes unless specified)	Duration of follow-up	Includes readmissions	Purpose of algorithm
Apte 2011	ICD-9: 998.5,998.51, 998.59	30d	Yes	Internal, comb
Hebden 2000	ICD-9 : 998.59	Unclear	Unclear	Internal, sole
Huang 2011	ICD-9: 34.01 34.02 34.10 86.01 86.04 86.09 86.22 86.28 91.71 91.72 91.73 513.1 519.2 682.2 682.3 682.8 686.8 686.9 730.00 730.08 730.09 730.20 730.28 730.29 730.30 730.38 730.39 730.80 730.88 730.39 730.90 730.98 730.99 785.52 790.7 875.0 879.8 879.9 891.0 891.1 996.60 996.61 99.62 996.71 998.31 998.32 998.51 998.83 998.9 CPT: 10060 10061 10140 10160 10180 11010 11040 11041 11042 11043 11044 12020 12021 13160 50000 50005 39000 39010; The algorithm was refined after piloting; unclear which codes are included in further analyses. Includes outpatient codes	60d	Yes	External
Platt 2002†	ICD-9: 998.0, 998.3, 998.5, 998.51, 998.59, 998.83, 780.6, 891.0, 891.1, 682.6, 682.9, 998.9 , 38.0 , 38.1, 38.10, 38.11, 38.19, 38.2, 38.3, 38.4, 38.40, 38.41, 38.42, 38.43, 38.44, 38.49, 38.8, 38.9, 790.7, 611.0, 682.0, 682.1, 682.2, 682.3, 682.4, 682.5, 682.6, 682.7, 682.8, 682.9, 686.0, 686.1, 686.8, 686.9, 958.3, 711.00, 996.6, 996.60, 996.61, 996.62, 996.63, 996.64, 996.65, 996.66, 996.67, 996.68, 996.69, 674.3, 879.0, 879.1, 879.2, 879.3, 879.4, 879.5, 879.6, 879.7, 879.8, 879.9, 875.0, 875.1 (also in outpatient setting). CPT: 87040, 87072, 87075, 87076, 87081, 87082, 87083, 87084, 10180, 11000, 11001, 15852 Note: the codes are included in a multivariable algorithm	30d	Yes	Internal, comb
Sands 2003†	Similar (or identical to Platt 2002)	30d	Yes	Internal, comb
Song 2008	ICD-9: 998.51, 998.59, 875.1, 519.2, 780.6	60d	Yes	Internal, comb
Yokoe 2004	ICD-9: 998.5, 998.51, 998.50	60d	Yes	Internal, comb

Abbreviations: CABG – coronary artery bypass graft, SSI – surgical site infection

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

SSI-Orthopedic

Study	Codes used (Inpatient only, primary & secondary unless specified)	Duration of follow-up	Includes readm	Purpose of algorithm
Bolon 2009	ICD-9: 998.5, 998.51, 998.51, 998.59, 996.66	365d	Yes	Internal, comb
Cadwallader, 2001	ICD-9: 996.66, 998.5, E878.1	30/365d	Yes	Internal, sole
Calderwood 2013	*THA:* ICD-9 Procedures: 84.56, 86.01, 86.22, 86.28 ICD-9: 686.8, 686.9, 711.00, 711.05, 711.08, 711.09, 711.40, 711.45, 711.48, 711.49, 711.90, 711.95, 711.98, 711.99, 730.00, 730.05, 730.08, 730.09, 730.10, 730.15, 730.18, 730.19, 730.20, 730.25, 730.28, 730.29, 730.90, 730.95, 730.98, 730.99, 996.60, 996.66, 996.67, 996.69, 998.5, 998.51, 998.59, 998.6 CPT: 10140, 10160, 10180, 12021, 13160, 20000, 20005, 26990, 26991, 26992, 27030, 27070, 27090, 27091, 27122, 27301, 27303, 35860 (includes outpatient)	365d	Yes	External
Inacio 2011	*1-120 day timeframe (wound only):* ICD-9: 998.30, 998.31, 998.32, 998.50, 998.51, 998.59, 680.5, 680.6, 680.9, 682.5, 682.6, 682.9, 686.9 *1-400 day timeframe (deep)* ICD 9: 711, 711.0, 711.00, 711.05, 711.06, 711.09, 711.60, 711.65, 711.66, 711.69, 711.90, 711.95, 711.96, 711.99, 730.00, 730.05, 730.06, 730.09, 730.20, 730.25, 730.26, 730.29, 730.90, 730.95, 730.96, 730.99, 996.6, 996.60, 996.66, 996.67, 999.3 ICD-9 Procedure: 80.00, 80.05, 80.06, 80.10, 80.16, 80.15, 78.60, 78.65, 78.66, 78.67, 78.69, 81.91, 86.04 (includes outpatient)	120d for superficial (wound) SSI 400d for deep SSI	Yes	Internal, sole

Abbreviations: SSI – surgical site infections, THA – total hip arthroplasty

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

SSI-other

Study	Target Procedure	Codes used (Inpatient only, primary & secondary unless specified)	Duration of follow-up	Includes readm	Purpose of algorithm
Campbell 2011†	Spinal surgery	Requested from corresponding authors; not available	LoS	No	Internal, sole
Daneman 2011	Caesarean section	ICD-10: O85002, O86002, O86004, O86009, O90202, K630, K750, L0331, L0332, L0333, N151, N730, K658, K650, O85004, N719, O86804, T813, T814, T857, T86842, T86822, T86882 (includes outpatient)	30	Yes	Internal, sole
Leth 2010†	Caesarean section	ICD-10: T81.4, O86.0 (incl. outpatient) Procedures: KLWB00, KMWB00, KLWC01, KMWC00, KMWC01	30	Yes	Internal, comb
Miner 2004†	Breast, caesarean section	*Caesarean section* ICD-9: 038 038.0 038.1 038.10 038.11 038.19 038.3 038.4 038.40 038.42 038.43 038.44 038.49 038.8 038.9 040.0 040.8 040.82 040.89 041 041.0 041.00 041.01 041.03 041.04 041.05 041.09 041.1 041.10 041.11 041.19 041.3 041.4 041.6 041.7 041.8 041.82 041.83 041.84 041.85 041.89 041.9 614.0 614.2 614.3 614.5 614.9 615 615.0 615.9 670 670.0 670.00 670.02 670.04 672 672.0 672.00 672.02 672.04 673.3 673.30 673.31 673.32 673.33, 673.34 682 682.2 682.5 686 686.8 686.9 780.6 790.7 996.6 996.60 996.62 996.69 998.5 998.51 998.59 Procedure: 86.01 86.04 86.22 10060 10061 10160 10180 11000 11001 *Breast* ICD-9: 675 675.0 675.00 675.01 675.02 675.03 675.04 675.1 675.10 675.11 675.12 675.13 675.14 675.2 675.20 675.21 675.22 675.23 675.24 675.8 (includes outpatient)	30/60	Yes	Internal, comb
Olsen 2010	Breast	ICD-9: 998.5, 998.51, 998.59, 996.69, 611.0, 682.2, 682.3 (in- and outpatient surgical care)	180	Yes	Internal, sole

Abbreviations: LoS – length of stay, SSI – surgical site infection

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

SSI – all/combined

Study	Procedure	Codes used (Inpatient only, primary & secondary unless specified)	Duration follow-up	Includes readm	Purpose of algorithm
Best 2002	All	ICD-9: 998.5	LoS	No	External
Calderwood, 2012	TKA, THA, Vascular surgery	**Limited list:** *TKA/THA:* ICD-9: 998.5, 998.51, 998.51, 998.59, 996.66 *Vascular:* ICD-9: 998.5, 998.51, 996.62 **Expanded list:** *THA:* Procedures: 84.56, 86.01, 86.22, 86.28 ICD-9: 686.8, 686.9, 711.00, 711.05, 711.08, 711.09, 711.40, 711.45, 711.48, 711.49, 711.90, 711.95, 711.98, 711.99, 730.00, 730.05, 730.08, 730.09, 730.10, 730.15, 730.18, 730.19, 730.20, 730.25, 730.28, 730.29, 730.90, 730.95, 730.98, 730.99, 996.60, 996.66, 996.67, 996.69, 998.5, 998.51, 998.59, 998.6 CPT: 10140, 10160, 10180, 12021, 13160, 20000, 20005, 26990, 26991, 26992, 27030, 27070, 27090, 27091, 27122, 27301, 27303, 35860 *TKA:* Procedures: 84.56, 86.01, 86.04, 86.22, 86.28 ICD-9: 686.8, 686.9, 711.00, 711.05, 711.06, 711.08, 711.09, 711.40, 711.45, 711.46, 711.48, 711.49, 711.90, 711.95, 711.96, 711.98, 711.99, 730.00, 730.05, 730.06, 730.08, 730.09, 730.10, 730.15, 730.16, 730.18, 730.19, 730.20, 730.25, 730.26, 730.28, 730.29, 730.90, 730.95, 730.96, 730.98, 730.99, 996.60, 996.66, 996.67, 996.69, 998.5, 998.51, 998.59, 998.6 CPT: 10140, 10160, 10180, 12021, 13160, 20000, 20005, 27301, 27303, 27310, 27488, 27603, 27604, 27607, 35860 *Vascular* Procedures: 54.0*, 54.19*, 86.01, 86.04, 86.22, 86.28 ICD-9: 686.8, 686.9, 996.6, 996.62, 998.51, 998.59, 998.6 CPT: 10140, 10160, 10180, 12021, 13160, 2000, 2005, 35840, 35840*, 35903, 35907* *only following a central vascular procedure (Includes outpatient codes)	Vasc: 60d TKA/ THA: 365d	Yes	Internal, sole
Curtis 2004	TKA, THA, vascular	ICD-10 AM mapped to Cadwallader et al (+ T84.41)	Unclear	Unclear	Internal, sole
Gerbier 2011	All	ICD-10: T814, T815, T816, T826, T827, T835, T836, T845, T846, T847, T857, O860 *refer to manuscript for extended selection	LoS	No	Internal, sole
Haley 2012†	CABG, colon, THA	ICD-9 : 5912, 567.21, 567.9, 682.2, 730.08, 730.25, 730.28, 995.91, 995.92, 996.66, 996.67, 996.77, 997.4, 998.11, 998.12, 998.30, 998.31, 998.32, 998.51, 998.59, 998.83, 38.11, 38.40, 41.09, 41.11, 41.12, 41.7, 41.85,	30/365	Yes	External
Hollenbeak 2011	General & vascular	ICD-9 : 998.59	30	Unclear	Internal, sole
Julian 2006	All	ICD-9: 730.09, 730.20-39, 730.90-730.99, 890.0-890.2, 891.0-891.2, 894.0-894.2, 996.61-996.63, 996.66, 996.67, 996.71, 996.72, 998.0, 998.31, 998.32, 998.51, 998.59, 998.6, 998.83, 999.3, 320.81, 320.82, 320.89, 320.0-320.3, 320.7, 320.9, 321.0-321.4, 321.8, 322.0, 322.9, 324.0, 324.1, 324.9, 420.90, 420.91, 420.99, 421.9, 422.90, 422.91, 513.1, 519.2, 682.1-682.4, 682.6, 682.7, 682.9, 728.0, 730.00-730.08 (PHC4 selection, secondary codes only)	LoS	No	External
Kanerva 2009	All	ICD-10 (first 3 slots): O86, T81.4, T84.5, T84.68, T82.7or A40, A41, A46, A48.8, A49, M00, M01, M46*B95.7 with or without T72.1, T21.2, Y83, Y84, Y88	LoS	No	Internal, sole

Study	Procedure	Codes used (Inpatient only, primary & secondary unless specified)	Duration follow-up	Includes readm	Purpose of algorithm
Lawson 2012	All	ICD-9: 998.5, 998.51, 998.59 Also includes outpatient	30	Yes	External
Lee 2011*	Gastric cancer patients	ICD-10 Mapped to PHC4 selection (see Julian)	Los	No	Internal, comb
Leth 2006†	Orthopedic Abdominal	ICD-10, T81.4	LoS	No	Internal, comb
Moro 2004	NNIS Procedures	ICD-9: three different sets of codes Group 1: 958.3, 996.60-996.69, 998.5, 998.51, 998.59 Additional group 2: group 1 + 254.1, 320.0, 320.2, 320.3, 320.8, 320.9, 321.0, 324.0, 324.1, 324.9, 2360.01, 360.00, 360.02, 360.04, 370.55, 373.13, 383.0-, 420.99, 421.0, 421.9, 424.90, 422.0, 422.90, 422.92, 422.99, 420.90, 447.6, 451-, 461.0-461.9, 475, 478.22, 478.24, 510.0-510.9, 513.0, 513.1, 519.2, 527.3, 528.3, 567.-, 566, 569.5, 572.0, 577.0, 590.10-590.11, 590.80, 590.2, 597.0, 597.80-, 599.0, 601.2, 604.0, 611.0, 614.0, 614.3, 614.5, 614.8, 614.9, 615.0, 615.9, 616.0, 616.1-, 675.10, 683, 711.0-, 711.4-, 711.6-, 711.8-, 711.9-, 727.00, 727.3,730.00-730.09.. Group 3: group 1 + group 2 + 998.6, 998.83, 999.3	LoS	No	Internal, comb
Sherman 2006*	All	ICD-9 as selected by PHC4 (see Julian)	LoS	No	External
Spolaore 2005	All	ICD-9: 998.5, 996.6 (not 996.64) or 958.3	LoS	No	Internal, comb
Stevenson 2008	All	Secondary ICD-9 as selected by PHC4 (see Julian). Outpatient codes unclear.	30/365	Yes	External
Tinelli 2011*	All	ICD-9 (up to 5 secondary): 264 codes, details not specified (no reply from corresponding author) Rehabilitation facility only 3x	LoS	No	Internal, sole
Verelst 2010	All	ICD-9: 998.51 or 998.59 in secondary diagnosis field, excl primary diagnoses for SSI and age < 16.	LoS	No	External
Yokoe 2012	Hysterectomy, vascular, colorectal	ICD-9: 998.5, 998.51, 998.59, 996.60, 996.62	30/365	Yes	Internal, comb

Abbreviations: CABG – coronary artery bypass graft, LoS – Length of Stay, PHC4 – Pennsylvania Healthcare Cost Containment Council, SSI – surgical site infection, THA – total hip arthroplasty, TKA – total knee arthroplasty,

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

CLABSI

Study	Denominator	Codes used (Inpatient only, primary & secondary unless specified)	Purpose of algorithm
Cevasco 2011	Within algorithm	PSI 7, version 3.1: ICD-9: 999.3, 999.62 in secondary diagnosis field; not PoA Excludes some high-risk patients based on primary diagnoses	External
Cima 2011	Within algorithm	Idem Cervasco 2011	External
Moehring, 2013	Within algorithm	CMS rule: 999.31 + PoA negative	External
Pokorny, 2006*	Unclear	ICD-9 codes for 'clinical infection: 038, 038.0, 038.1, 038.2, 038.3, 038.4, 038.8, 038.9, 360.0, 360.1, 480, 481, 482.0, 482.1, 482.2, 482.4, 482.8, 482.9, 483, 484, 485, 486, 590.10, 595.0, 599.0, 646.60, 646.61, 646.62, 646.63, 646.64, 646.6[0-4], 670, 670.02, 670, 674.34 [4], 790.7, 421.0 , 421.1, 421.9, 996.6, 996.61, 996.62, 996.64, 996.69, 998.5, 998.51, 998.59	Internal, comb
Scanlon 2008	Within algorithm	Pediatric quality indicator: 999.3, 999.62 (does not include PoA indicator) Denominator: Age 0 – 17, admitted without infection as primary diagnosis,	External
Sherman 2006*	Within algorithm	ICD-9: specified by PHC4 (secondary diagnoses) 0380, 038.1, 038.11, 038.19, 038.2, 038.3, 38.40, 38.41, 38.42, 38.43, 38.44, 38.49, 38.8, 38.9, 790.7, 995.9, 995.91, 995.92, 995.92	External
Stamm 2012	Identified by traditional surveillance	ICD-9; details not specified (no reply from corresponding author)	Internal, sole
Stevenson 2008	Patients with a positive blood culture	ICD-9: specified by PHC4 (secondary diagnoses) 0380, 038.1, 038.11, 038.19, 038.2, 038.3, 38.40, 38.41, 38.42, 38.43, 38.44, 38.49, 38.8, 38.9, 790.7, 995.9, 995.91, 995.92, 995.92	External
Stone 2007	Within algorithm	PSI 7, version 2.1	External
Tehrani 2013	Sens: patients in routine surveillance PPV: within code selection	CMS HAC rule: 999.31 + PoA negative	External
Zrelak 2011	Within algorithm	PSI 7, version 3.1: ICD-9: 999.3, 999.62 in secondary diagnosis field; not PoA Excludes some high-risk patients from denominator based on primary diagnoses	External

Abbreviations: CLABSI – central-line associated bloodstream infection, CMS – Centers for Medicare and Medicaid Services, HAC – Hospital-acquired condition, PoA – present on Admission, PHC4 – Pennsylvania Health Care Cost Containment Concil, PSI – patient safety indicator,

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

Bloodstream infection/Sepsis

Study	Codes used (Inpatient only, Primary & secondary unless specified)	Purpose of algorithm
Best 2002	ICD-9: 998.0 - 38.0 - 38.9, 785.5, 785.59	External
Braun 2006†	Compares several algorithms at the aggregate level. Does not detail all algorithms	External
Cevasco 2011a	PSI 13, version 3.1 Secondary ICD9 diagnoses (not PoA) : 038.0, 38.1, 038.10, 38.11, 038.12, 38.19, 38.2, 0383, 785.52, 785.59, 998.0, 995.91, 995.92, 038.4, 038.41, 038.42, 038.43, 038.44, 038.49, 038.8, 0389. Numerator: Patients aged over 18 undergoing an elective procedure with LoS > 3 days . Excludes patients with principal diagnosis of infection/sepsis, patients with infection PoA, patients with cancer/immunosuppression and obstetric admissions.	External
Cevasco 2011b	PSI 13, version 3.1 (idem Cevasco 2011a)	External
Cima 2011	PSI 13, version 3.1 (idem Cevasco 2011a)	External
Gerbier 2011	ICD-10: A021, A207, A217, A227, A241, A267, A280, A327, A392, A393, A394, A40-, A41-, A427, A483, A499,A548, B007, B377, O080, O753, O85, P3600, P3610, P3620,P3630, P3640, P3650, P3680, P3690	Internal, sole
Hougland 2008	ICD-9: 038.0, 038.10, 038.11, 038.19, 038.3, 038.40, 038.41, 038.42, 038.43, 038.44, 038.49, 038.8, 038.9, 790.7	Ext
Kanerva 2009	ICD-10 (first 3 slots): A40, A41, B37, R 50.9, J15.9, J 18.9, K80, N30 with or without Y82, Y83	Internal, sole
Koch 2012	PSI 13, version 4.2 Secondary ICD9 diagnoses (not PoA) : 038.0, 38.1, 038.10, 38.11, 038.12, 38.19, 38.2, 0383, 785.52, 785.59, 998.0, 995.91, 995.92, 038.4, 038.41, 038.42, 038.43,038..44, 038.49, 038.8, 0389. Numerator: Patients aged over 18 undergoing an elective procedure with LoS > 3 days . Excludes patients with principal diagnosis of infection/sepsis, with infection PoA, with cancer/immunosuppression and obstetric admissions.	Ext
Lawson 2012	ICD-9: 038*, 785.52, 995.91, 995.92, 998.0, 998.59, 999.31 (incl outpatient)	External
Lee 2011*	ICD-10 Mapped to PHC4 selection: 0380, 038.1, 038.11, 038.19, 038.2, 038.3, 38.40, 38.41, 38.42, 38.43, 38.44, 38.49, 38.8, 38.9, 790.7, 995.9, 995.91, 995.92, 995.92. No reply from corresponding author regarding exact code selection.	Internal, comb
Murff 2011	PSI 13, version 3.1	External
Ollendorf 2002	Presence of codes indicative of sepsis on first 9 positions of UB-92 bill 003.1, 020.2, 022.3, 036.2, 038.0 038.1, 038.2, 038.3, 038.4, 038.41, 038.42, 038.43, 038.44, 038.49, 038.8, 038.9, 054.5, 790.7,	Internal, sole
Romano 2009	PSI 13 version 2.1 (ICD-9). Original: any 38.xx code in secondary diagnosis field. Revised: 38.xx code in secondary diagnosis field or code 998.0, 998.1, 785.59, 785.5, 785.52 No accounting for PoA. Denominator same as other PSI studies	External
Scanlon 2008	PDI (ICD-9). Numerator: secondary diagnosis code for sepsis, without PoA indicator Denominator: Age 0-17, non-neonate, LoS > 4 days, without sepsis of infection as primary diagnosis	External
Verelst 2010	PSI 13, version 3.1 (see Cevasco 2011a)	External

Abbreviations: CLABSI – central-line associated bloodstream infection, CMS – Centers for Medicare and Medicaid Services, HAC – Hospital-acquired condition, LoS – length of stay, PoA – present on Admission, PHC4 – Pennsylvania Health Care Cost Containment Concil, PDI – pediatric quality indicator, PSI – patient safety indicator,

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

CAUTI

Study	Denominator	Codes used (Inpatient only, primary & secondary unless specified)	Purpose of algorithm
Meddings 2010	Within algorithm (996.64)	ICD-9: Secondary code 112.2, 590.10, 590.11, 590.2, 590.3, 590.80, 590.81, 595.0, 597.0, and 599.0 with or without PoA.	External
Pokorny 2006*	Unclear	ICD-9 codes for 'clinical infection, see under CLABSI	Internal, comb
Sherman 2006*	Within algorithm	ICD-9: 590.00, 590.01, 590.1, 590.11, 590.2, 590.3, 590.8, 590.9, 595.0, 595. 1, 595.2, 595.3, 595.81, 595.89, 595.9, 599.0, 9975.	External
Zhan 2009	Within algorithm 1. Procedure code 57.94 or 57.95 2. Claims with major surgery 3. Claims with any ICD-9 procedure code	ICD-9 in secondary diagnosis fields: 996.64, 112.2, 590.10, 590.11, 590.2, 590.8, 590.81, 590.9, 595.0, 595.3, 595.4, 595.89, 595.9, 597.0, 597.80, 599.0 Excluding discharges with primary discharge codes for sepsis or infection or any discharge code for immunosuppression (in analogy to PSI)	External

Abbreviations: CAUTI – catheter-associated urinary tract infection, PoA – present on admission, PSI – patient safety indicator

UTI

Study	Codes used (Inpatient only, primary & secondary unless specified)	Purpose of algorithm
Best 2002	ICD-9: 599.0, 590.1 - 590.9, 595.0 - 595.9	External
Campbell 2011†	Requested from corresponding authors; not available	Internal, sole
Gerbier 2011	ICD-10: N300, N34-, N390, O862, O863, T835	Internal, sole
Heisler 2009	Hospital adaptation of ICD-9 codes, equivalent to 599.0 and 999.64	Internal, sole
Julian 2006	ICD-9: 590.00, 590.01, 590.10, 590.11, 590.2, 590.3, 590.80, 590.9, 595.0-595.3, 595.81, 595.89, 595.9, 599.0, 997.5 (secondary codes only, PHC4)	External
Kanerva 2009	ICD-10: N30, N39, A41, R50.9; first three slots only	Internal, sole
Landers 2010	ICD-9: 599.0	Internal, sole
Lawson 2010	ICD-9: 112.2, 590.1*, 590.3, 590.8*, 595.0, 595.30, 599.0, 996.64	External
Lee 2011*	ICD-10 Mapped to PHC4 selection (see Julian) No reply from corresponding author regarding exact code selection.	Internal, comb
Tinelli 2011*	ICD-9 (up to 5 secondary): 264 codes, details not specified (no reply from corresponding author) Rehabilitation facility only	Internal, sole

Abbreviations: UTI –urinary tract infection, PHC4 – Pennsylvania Health Care Cost Containment Council.

Table S6: Administrative data algorithm, by HAI type, cont'd

*: studies presenting only accuracy estimates aggregated over multiple types of HAI.

†: studies assessing an algorithm combining multiple sources of data that did not allow for data-extraction for administrative data only.

VAP

Study	Denominator	Codes used (Inpatient only, primary & secondary unless specified)	Purpose of algorithm
Drees 2010	Within algorithm	ICD-9: 999.9	Internal, sole
Julian 2006	Within algorithm (code for mechanical ventilation)	ICD-9 (secondary codes only according to PHC4): 480.0-480.3, 480.8, 480.9, 481, 482.0-482.2, 482.30-482.32, 482.39-482.41, 482.82-482.84, 482.89, 482.9, 483.0, 483.1, 483.8, 485, 486, 487.0, 482.49, 482.81	External
Pokorny 2006*	Unclear	ICD-9 codes for 'clinical infection, see under CLABSI	Internal, comb
Sherman 2006*	Within algorithm	ICD-9 (secondary codes only according to PHC4): 480.0-480.3, 480.8, 480.9, 481, 482.0-482.2, 482.30-482.32, 482.39-482.41, 482.82-482.84, 482.89, 482.9, 483.0, 483.1, 483.8, 485, 486, 487.0, 482.49, 482.81	External
Stamm 2012	Identified by traditional surveillance	ICD-9; details not specified (no reply from corresponding author)	Internal, sole
Stevenson 2008	Patients with ventilator procedure code (31.1, 31.2, 31.29, 31.21, 96.04, 96.7, 96.70, 96.71, 96.72)	ICD-9 (secondary codes only according to PHC4): 480.0-480.3, 480.8, 480.9, 481, 482.0-482.2, 482.30-482.32, 482.39-482.41, 482.82-482.84, 482.89, 482.9, 483.0, 483.1, 483.8, 485, 486, 487.0, 482.49, 482.81	External
Verelst 2010	Belgian nomenclature code for artificial ventilation (211046)	PSI version 3.1 ICD-9 codes for pneumonia in secondary field. Excludes primary diagnosis of pneumonia or 997.3, or viral pneumonia, immunocompromised, < 16 years.	External

Abbreviations: PHC4 – Pennsylvania Health Care Cost Containment Council, PSI – patient safety indicator, VAP – ventilator-associated pneumonia.

Pneumonia (sometimes also including VAP)

Study	Codes used (Inpatient only, primary & secondary unless specified)	Purpose of algorithm
Best 2002	ICD-9: 997.3, 480.0 - 487.0	External
Gerbier 2011	ICD-10: J10-, J11-, J12-, J13-, J14-, J15-, J16-, J17- , J18-,	Internal, sole
Hougland 2008	ICD-9: 481, 482.0, 482.1, 482.2, 482.30, 482.31, 482.32, 482.39, 482.40, 482.41, 482.49, 482.81, 482.82, 482.83, 482.84, 482.89, 482.9, 483.8, 485, 486.	External
Kanerva 2009	ICD-10: J13, J15.9, J18.9, J20.9, J60.9, J05, J38.5, B59, R91; first three slots only	Internal, sole
Lawson 2012	ICD-9: 39.1, 1124, 1179, 1363, 4466.19, 480*, 481, 482*, 483*, 4841, 4846, 4847, 485, 486, 4870, 507*, 5130, 5168, 997.31, 997.39	External
Lee 2011*	ICD-10 Mapped to PHC4 selection (see Julian). No reply from corresponding author regarding exact code selection.	Internal, comb
Murff 2011	PSI version 3.1 for pneumonia as a component of *Failure to Rescue (PSI 4)* ICD-9 codes: 482.0, 482.1, 482.2, 482.3, 482.3, 482.30, 482.31, 482.32, 482.39, 482.4, 482.40, 482.41, 482.49, 482.8, 482.81, 482.82, 482.83, 482.84, 482.89, 482.9, 485, 486, 507.0 514, excluding cases with a pre-existing condition of pneumonia or 997.3, with any diagnosis code for viral pneumonia, MDC 4 (diseases/disorders of respiratory system) or with any diagnosis of immunocompromised state In this study, the PSI patient population was limited to patients eligible for both the VASQIP measures and PSI criteria (see the article for details).	External

Abbreviations: PHC4 – Pennsylvania Health Care Cost Containment Council, PSI – patient safety indicator, VAP – ventilator-associated pneumonia.

Other

Study	Target infection	Codes used (Inpatient only, primary & secondary unless specified)	Purpose of algorithm
Van Mourik 2013	Drain-related meningitis	ICD-9: 112.83, 320.00 – 320.9, 322.00 – 322.9, 324.00 – 324.9, 349.10, 792.00, 996.60, 996.63, 996.70, 996.75, 997.00, 997.01, 997.09, 998.50 – 998.59, 999.30 – 999.39 Patients at risk identified by manual surveillance	Internal, sole
Yokoe 2001†	Post-partum infection	ICD9: 670.2, 670.04, 599.0, 674.34, 675.14, 675.24, 998.5 COSTAR (ambulatory): DA140, DC150, DC408, DH140, DL101, DM153, DR180	Internal, comb

Abbreviations: COSTAR: Computer-stored ambulatory record.