Healthcare-Associated Infections: making sensible sense of electronic record data

Maaike van Mourik

UMC Utrecht

- Understand the basic principles of semi- and fully-automated surveillance
- Having a general impression of the data sources needed for automated surveillance
- Grasping the importance of clinical context when developing automated surveillance methods
- Understand the consequences of automated surveillance w.r.t. interpretation of surveillance outcomes.

Topics

- Surveillance: Why and how?
- Why automated surveillance?
- Some terminology
- Semi-or fully automated surveillance
- Commonly used data sources
- Algorithms
- Shifting definitions?
- Risks and limitations

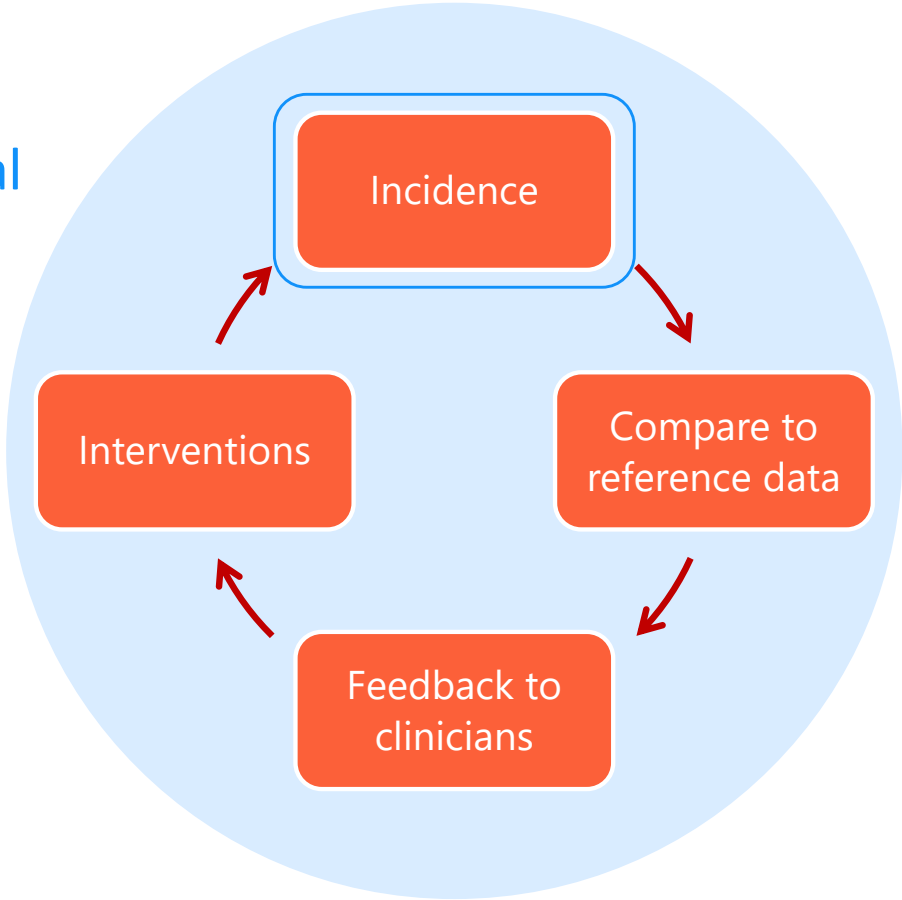
Surveillance of HAI

"systematic collection, analysis, interpretation and dissemination of data regarding a health-event for use in public health action to reduce morbidity and mortality and to improve health"

- SSI, CLABSI, UTI…
- 1 in 25 patients admitted to hospital

Surveillance:

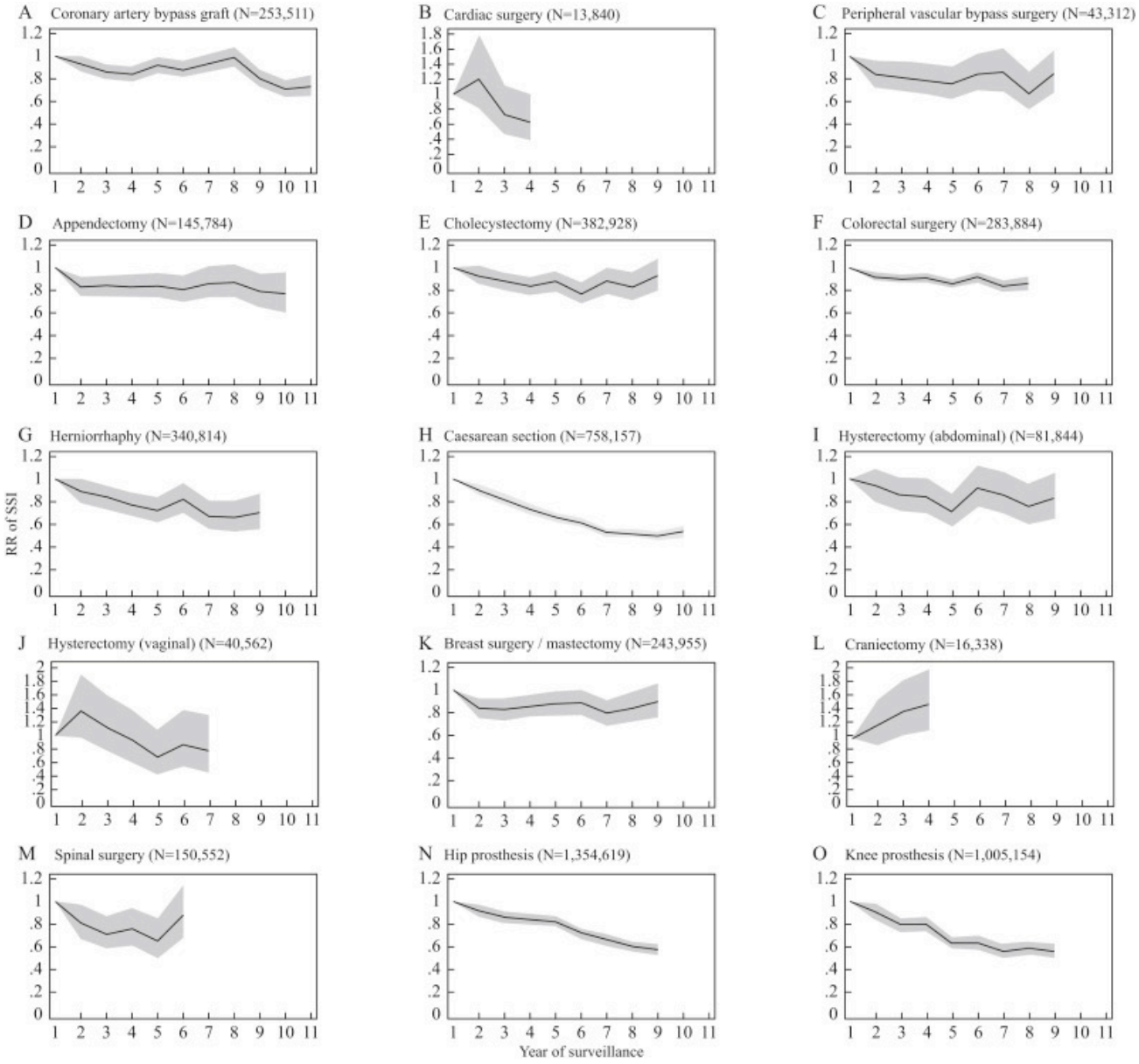
- Within 1 facility
- National networks (PREZIES, KISS)
- Mandatory or voluntary participation
- Confidential or public data



UMC Utrecht

Haley, 1985, World Health Organization 2011, Umscheid *ICHE* 2014, Magill NEJM 2018

Is surveillance useful?



Surveillance
=
data 4 action

UMC Utrecht

Abbas et al JHI 2019, Haley 1985,

Surveillance

Conventional surveillance

- Manual, retrospective chart review
- Determine infection status based on case-definition
- Data collection incl risk factors
- Reports & interpretation

- *Labour-intensive*
- *Prone to error*
- *"The more you look, the more you find"*

Why automated surveillance?

- More efficient by reducing workload
- Better standardization
- Less subjective interpretation
- Less effort-dependent

Haley *Am J Epidemiol.* 1985, Talbot et al. *Ann Intern Med.* 2013. Trick *Clin Infect Dis.* 2013, De Bruin *JAMIA* 2014, Freeman *J Hosp Infect* 2014

Terminology

Automated surveillance (AS) – Any form of surveillance where (parts of) the manual assessment are **replaced by an automated process**. This includes fully automated and semi-automated detection of HAI and collection, validation and analysis of denominator data. AS is based on routine care data, usually by applying appropriate algorithms.

Routine care data – All data documented in an electronic format during the routine process of care, for example surgical procedures, prescriptions and diagnostic testing results. These data may be stored and accessed in various IT systems.

Source data – (Raw) data elements from routine care data used by algorithms to detect (possible) HAI, calculate the denominator or risk factors. Examples include microbiology results, admission and discharge dates, central line days, procedure codes.

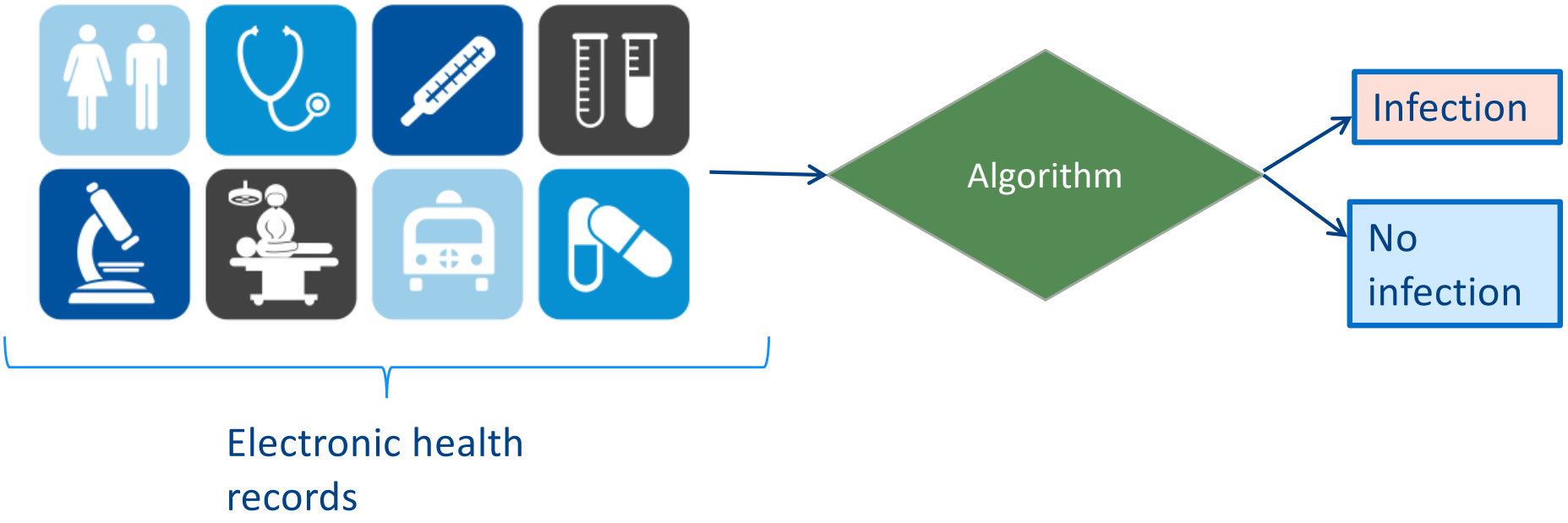
HAI surveillance result – Individual-level HAI status data (HAI yes or no, including details of HAI) and denominator data (e.g. central line days, surgical procedures).

Observed HAI rate – Aggregate crude rate of HAI calculated based on HAI surveillance result, e.g. incidence density rate.

Automated surveillance

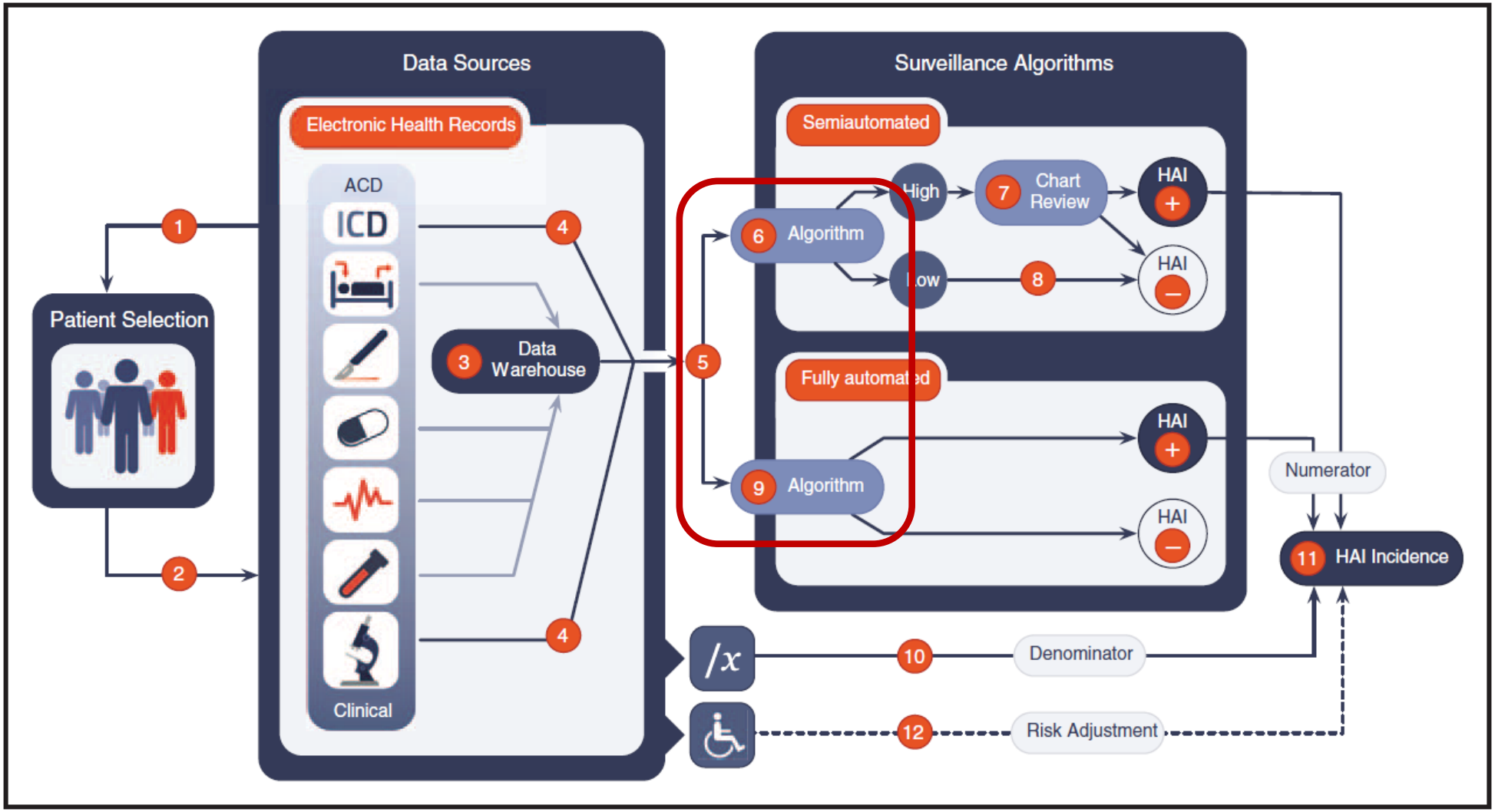
Does not mean: electronic **documentation** of infections in electronic health records

It does mean: re-using data from electronic health records to **take decisions**.



UMC Utrecht

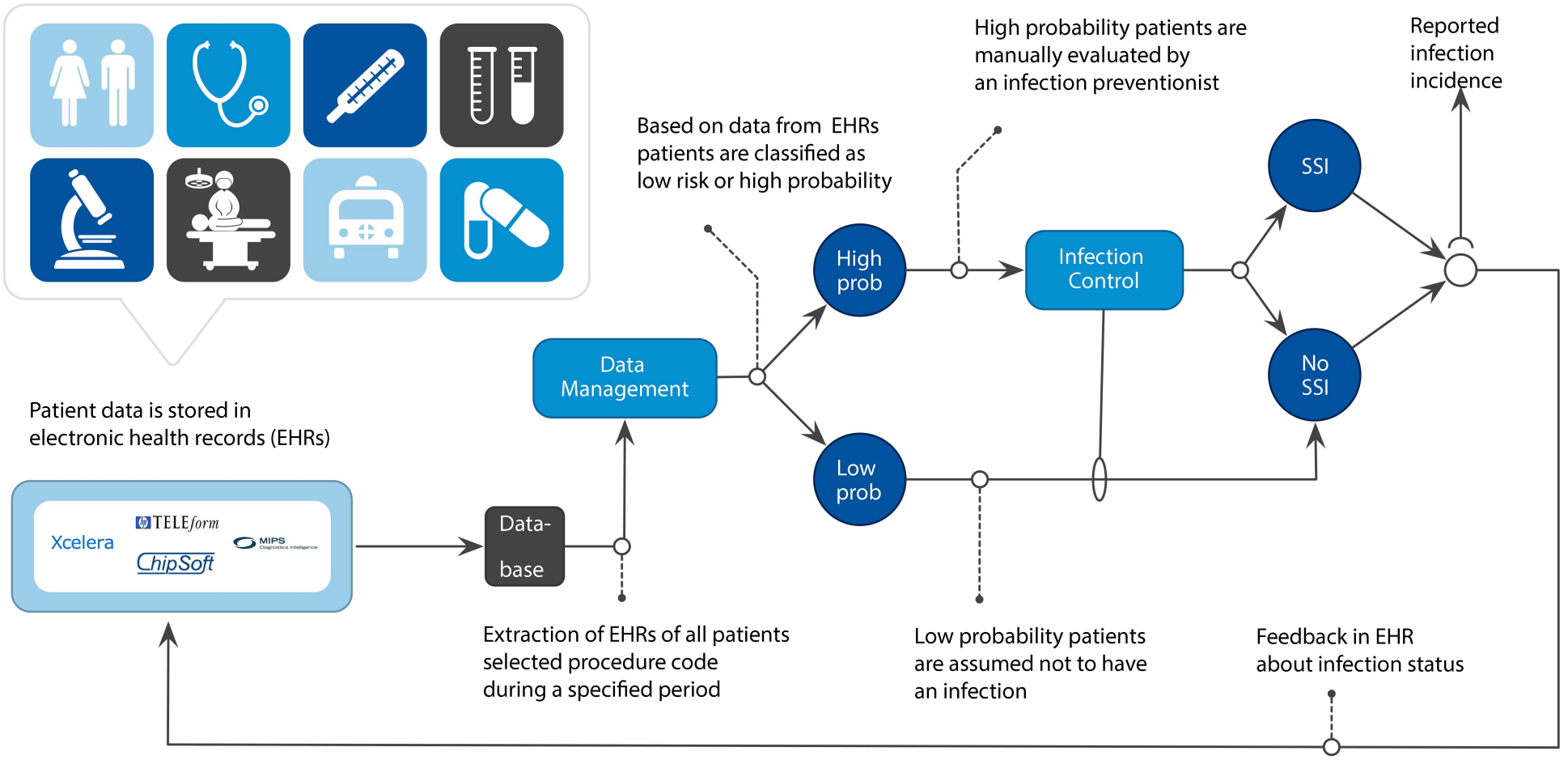
The bigger picture



Sips et al *Curr Topics Infect Dis* 2017

Types of automated surveillance

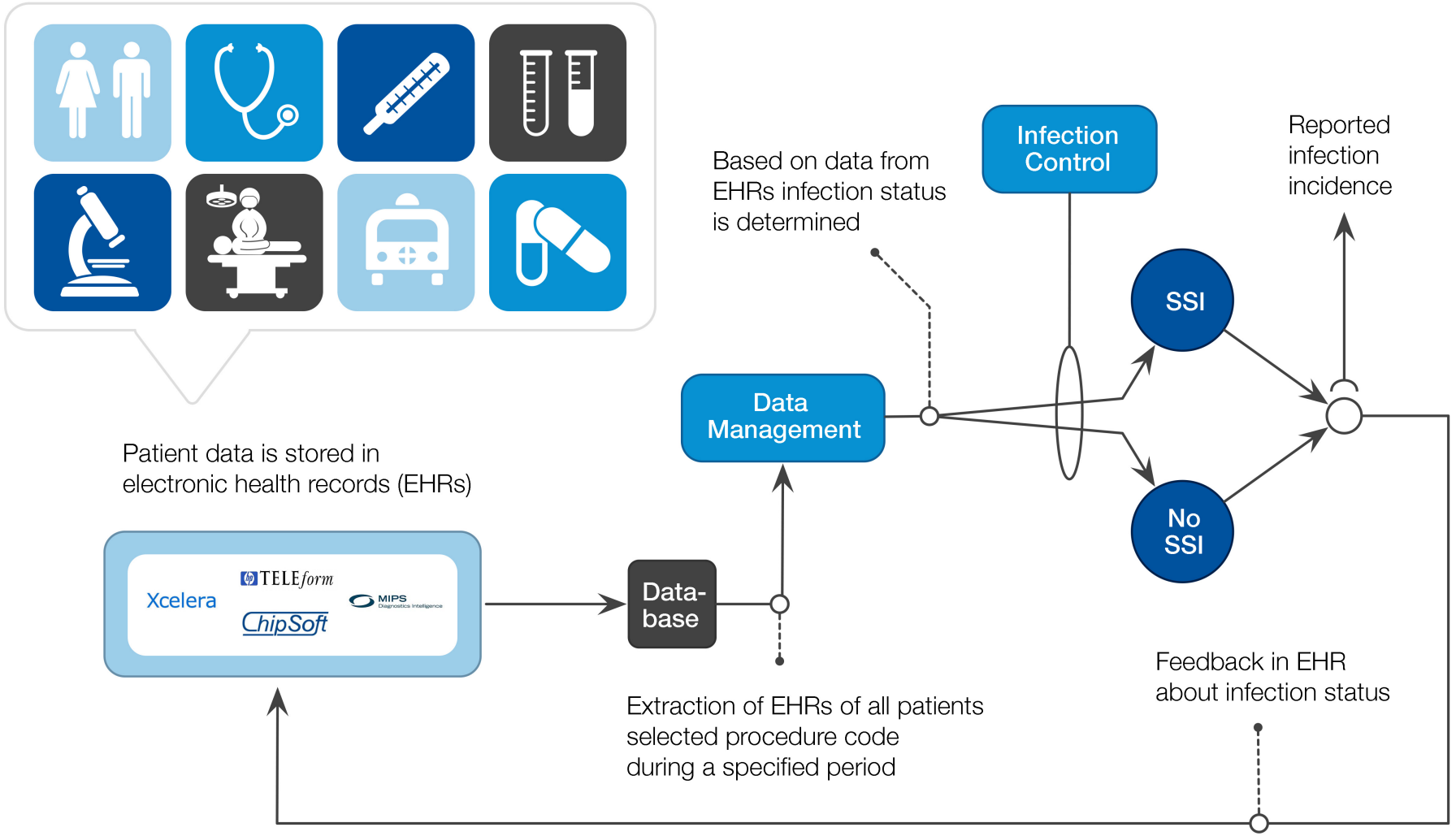
- **Semi-automated:** Select possible cases of infection for manual confirmation by chart review.
  - Aim to find all possible cases (sensitivity)



Woeltje *J Hosp Infect.* 2013. Figure by Meander Sips

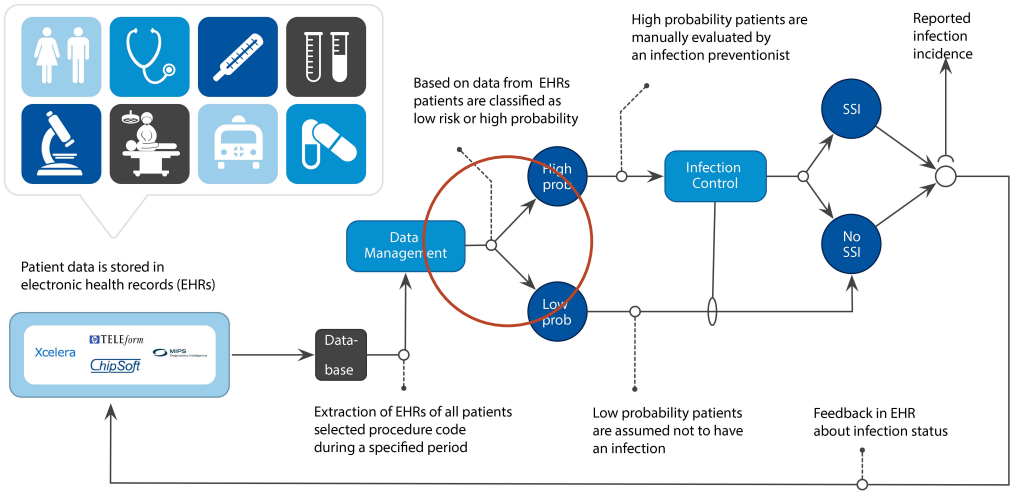
Types of automated surveillance

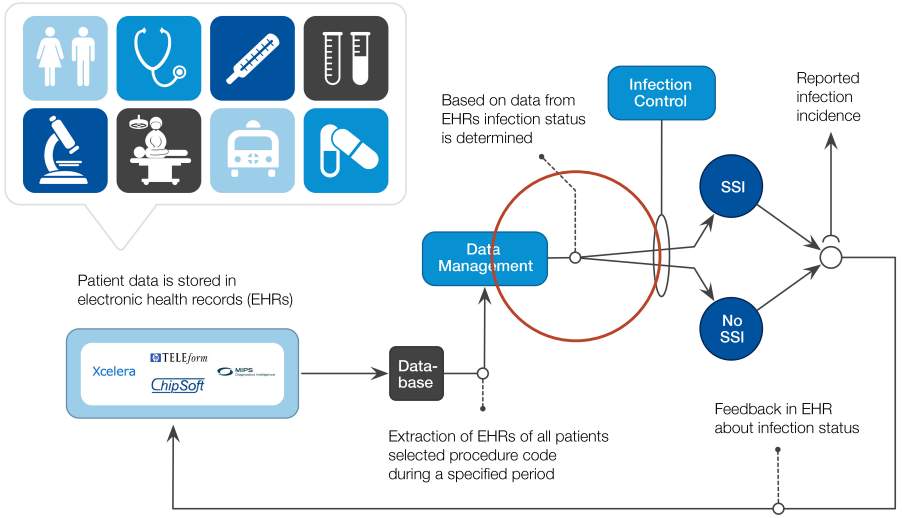
- **Fully automated:** No manual confirmation of infections
  - Direct comparison of rates -> comparability



Woeltje *J Hosp Infect.* 2013. Figure by Meander Sips

Examples





UMC Utrecht

SSI after hip or knee replacement

Reoperation by same specialty

> 5 culture relevant taken OR positive culture

Prolonged antibiotics (>14 days)

Readmission or prolonged LOS

Classification algorithm
≥3 out of 4

High probability, chart review

Low probability, no SSI

SSI

No SSI

Performance	Sensitivity (%)	% workload reduction
Hip/knee	100	95

Sips et al 2017

TABLE 1. Performance of Individual Predictors, Diagnostic Categories, and Models[a]

Variable	Deep SSI		Chart Review			
	Yes (n = 30)	No (n = 1,607)	No. (n = 1,637)	%	Sensitivity, %	PPV, %
Case-finding in routine surveillance						
≥1 relevant microbiological culture obtained	30	358	388	23.7	100.0	7.7
Diagnostic category 1: Microbiology						
1A ≥1 positive relevant culture	30	81	111	6.8	100.0	27.0
1B ≥5 relevant cultures obtained	30	58	88	5.4	100.0	34.1
1 Total: 1A or 1B	30	111	141	8.6	100.0	21.3
Diagnostic category 2: Antibiotics						
2 ≥14 d of antibiotic exposure	30	50	80	4.9	100.0	37.5
Diagnostic category 3: (Re)admissions						
3A Primary admission ≥14 d	16	220	236	14.4	53.3	6.8
3B ≥1 readmission for a relevant specialty	23	90	113	6.9	76.7	20.4
3 Total: 3A or 3B	30	295	325	19.9	100.0	9.2
Diagnostic category 4: Surgery						
4 ≥1 orthopedic surgical procedure	30	90	120	7.3	100.0	25.0
Surveillance models						
m_4 Positive on 4 categories	30	14	44	2.7	100.0	68.2
m_3 Positive on ≥3 categories	30	46	76	4.6	100.0	39.5
m_2 Positive on ≥2 categories	30	128	158	9.7	100.0	19.0
m_1 Positive on ≥1 category	30	358	388	23.7	100.0	7.7

NOTE. SSI, surgical site infection; PPV, positive predictive value.

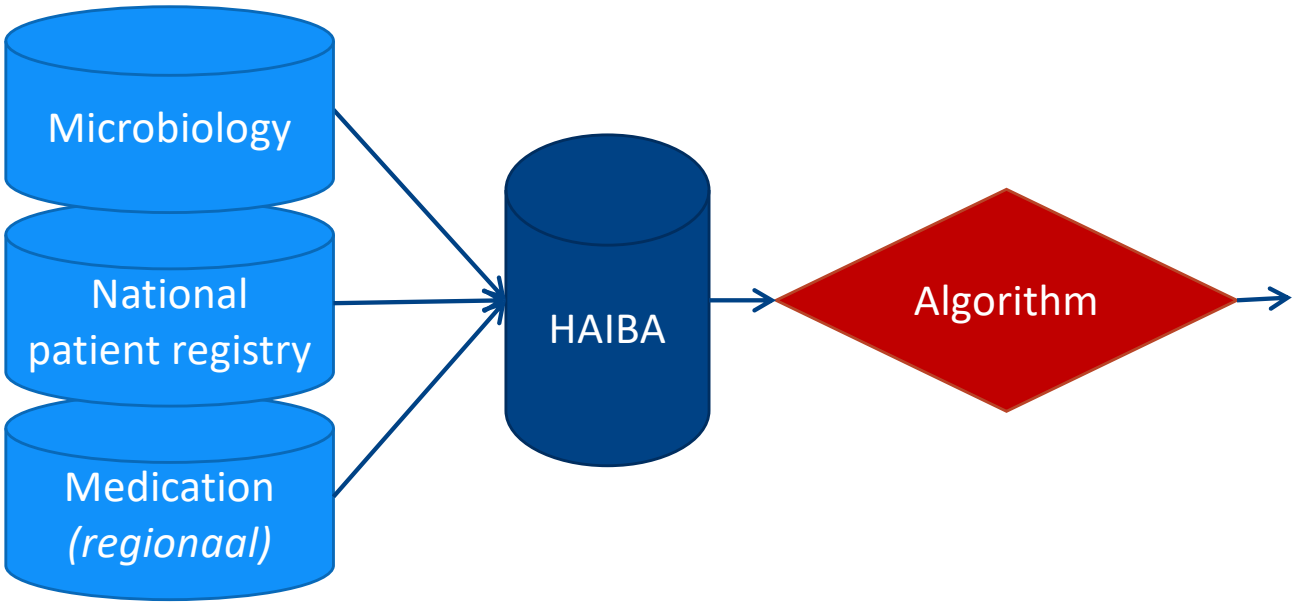
[a]All predictors were analyzed for a period of 120 days after the primary procedure (Appendix 1 for details). The surveillance models include 4 main diagnostic categories (1–4), of which 2 are further subdivided (A and B). Only 1 subcategory needs to be fulfilled for the main diagnostic category to turn out positive. The proposed models are designated as m_i, in which *i* represents the minimum number of main diagnostic categories a patient needs to fulfill to be selected for chart review.

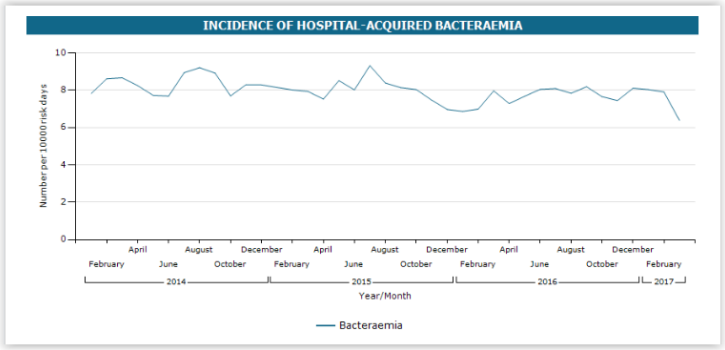
Multicenter validation

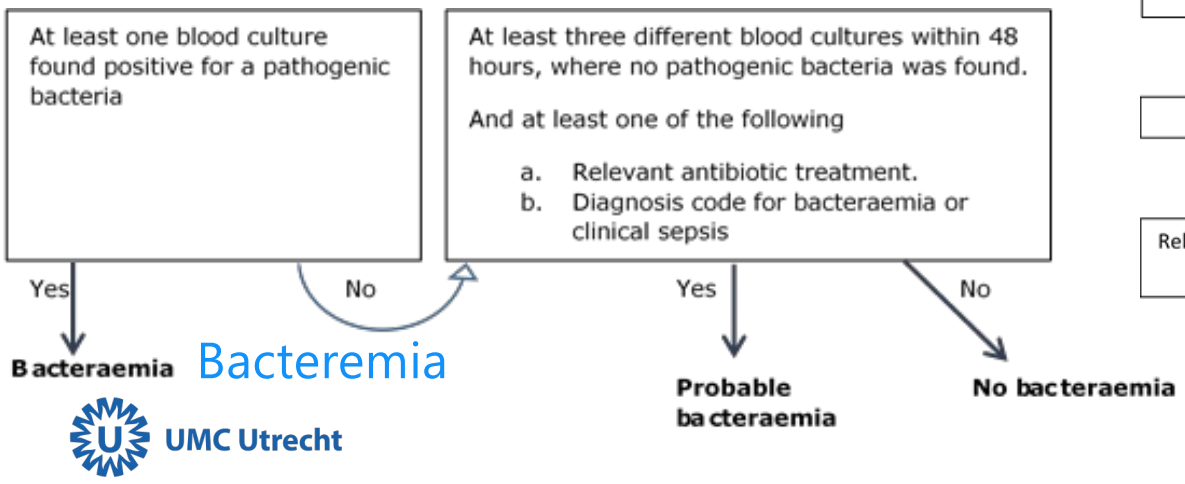
	Sensitivity, % (95%CI)	PPV, % (95%CI)	Work load reduction%
Hospital A	100 (86.6-100)	72.2 (54.8-85.8)	98.5
Hospital B	95.7 (78.0-99.9)	68.8 (40.0-83.3)	98.0
Hospital C	100 (78.2-100)	57.7 (36.9-76.7)	98.5
Hospital D	93.6 (78.6-99.2)	55.8 (41.3-69.5)	98.4

Verberk, 2020

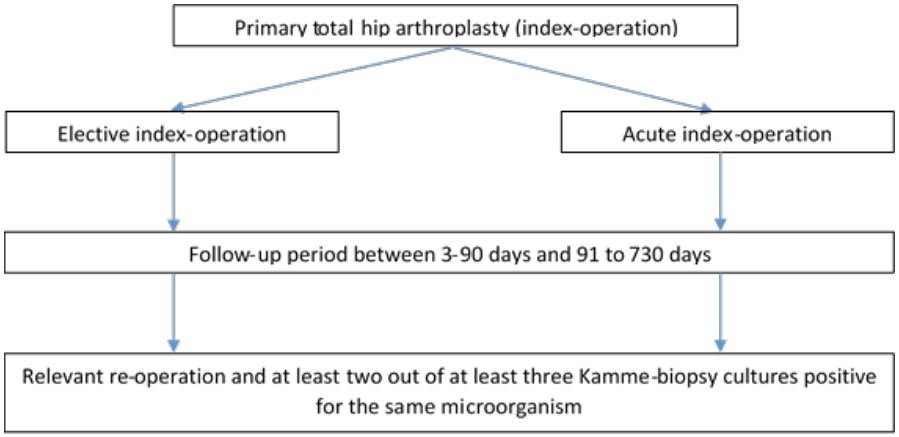
National automated surveillance (HAIBA)







Bacteremia



SSI after THP/TKP

UMC Utrecht

http://www.esundhed.dk/sundhedskvalitet/HAIBA/Sider/Documentation.aspx

Many many ways to get there! But how to do it?

Rationale PRAISE network

Approaches	Semi- or fully automated? Adapted definitions?
Data sources	Clinical or administrative? Structured or unstructured
Organization	Infrastructure? Responsibilities

Initiated in 2019

Heterogeneity in automated surveillance methods

Stand-alone development is inefficient

Many shared barriers and challenges
Inefficient use of resources
Risk losing comparability

Providing a **R**oadmap for **A**utomated **I**nfection **S**urveillance in **E**urope.

UMC Utrecht

Aim of PRAISE network

Provide guidance on how to move automated surveillance from research setting to large-scale implementation

- High-level conceptual guidance
- Address IT and Governance aspects in accompanying papers
- Hospitals & surveillance networks can translate to their local setting to support design and implementation

full supplement available online

SCAN ME

UMC Utrecht

Published as a supplement in *Clinical Microbiology & Infection. Vol. 27 S1*
https://www.sciencedirect.com/journal/clinical-microbiology-and-infection/vol/27/suppl/S1

Selected topics

- **Semi or fully-automated surveillance**
- **Data sources**
- **Centrally or locally implemented surveillance**
- **Choosing your algorithms**
- **Shifting definitions**
- **Risks of automated surveillance**

1. Semi- vs. fully automated surveillance

	Semi-automated	Fully automated
Chart review?	Selected cases	None
Performance	1. Sensitivity 2. Workload reduction	1. Specificity
Data requirements	Standardised data	Standardised data
Case-definition	Standardised definition	Adapted definition (indicator)
Subjectivity	Partial, some chart review required (advantage?)	No room for subjective interpretation
Acceptance	Clinical buy-in	Clinical buy-in less certain
	…	…

2. Data sources

Routine care data:

- collected during routine process of care
- stored in EHR
- extracted through clinical data warehouses

→

- ✓ Availability in a standardized format differs
- ✓ Depends on clinical practice and documentation
- ✓ Additional registration burden?

Exact requirements depend on target of surveillance

Clinical data	Medico-administrative
Microbiology results	**Medication use**
Laboratory results	**Procedure codes**
Device use	**Diagnosis codes**
Physician narratives*	**Billing data**
Other diagnostics (radiology)*	

*often free text

Table 2
Categories to indicate the suitability of surveillance data in a hospital usable for

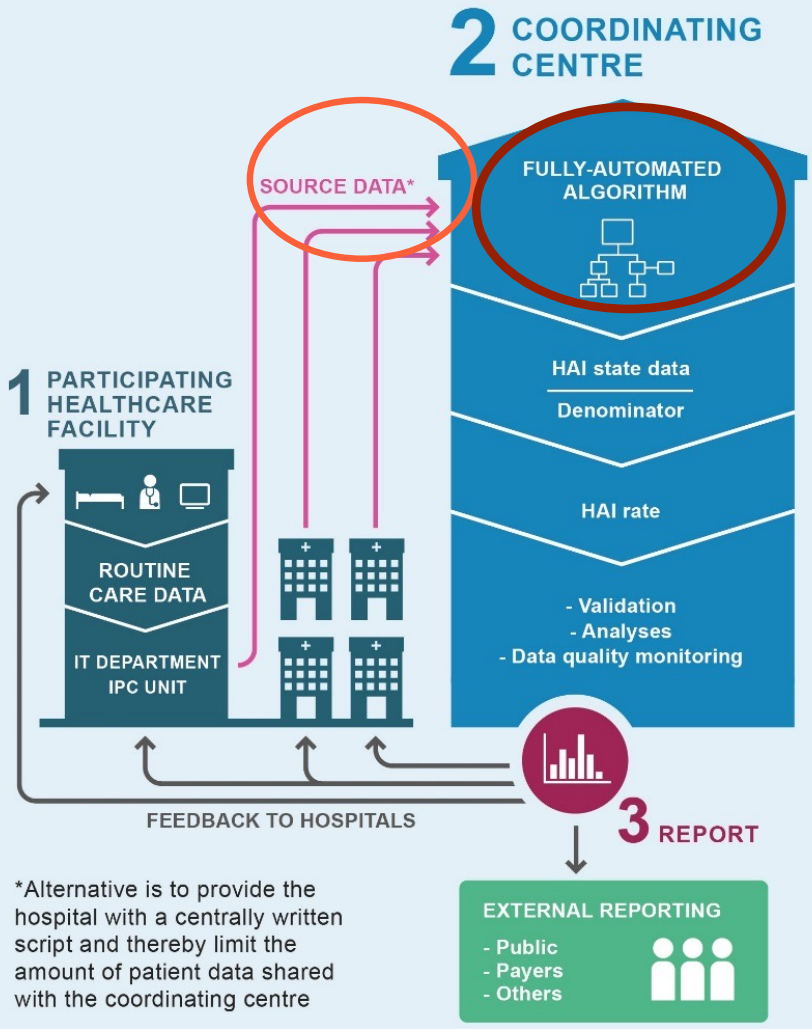
Surveillance data		2	3	4	5
Data already exist in a digital subsystem		Yes	Yes	Yes	No
Data are structured and well defined	Yes	Yes	Yes	No	No
Data are available in most facilities and sem	Yes	Yes	No	No	No
Data are accessible for surveillance algorithm	Yes	No	No	No	No

Quality Quality Quality! Validate Validate Validate

De Bruin *JAMIA 2014*, van Mourik BMJ Open 2015, Behnke *CMI* 2021

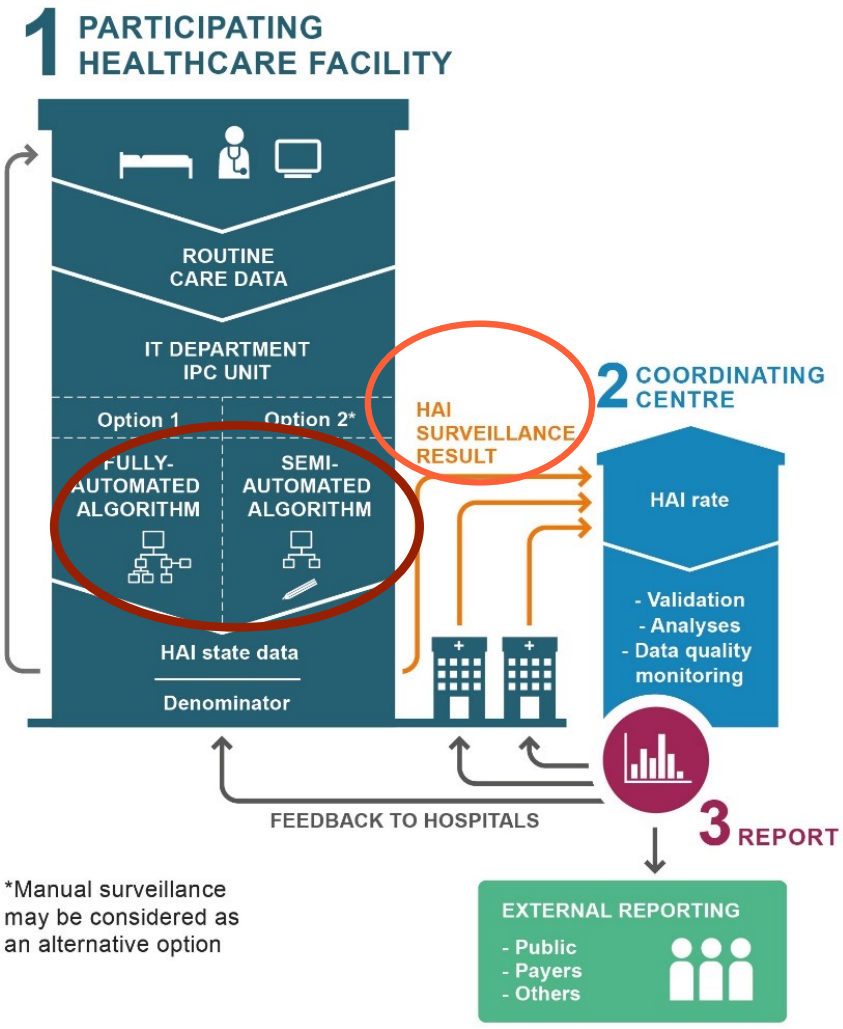
3. Surveillance in network: local or central?

Centrally Implemented Surveillance



*Alternative is to provide the hospital with a centrally written script and thereby limit the amount of patient data shared with the coordinating centre

Locally Implemented Surveillance



*Manual surveillance may be considered as an alternative option

Local	Centralized
Adapt to local IT infrastructure	Enforce fixed infrastructure
Custom-built methods for situation	Standardized methods
Shared specifications?	Shared specifications required
More limited local knowledge	Centralized knowledge
…	…

4. Choosing your algorithm (semi-automated)

Study the literature or develop your own

Align algorithm with clinical practice
- Do not over-specify & allow room for practice variation

Perform (retrospective) validation
- Source data
- Algorithm classification
- Risk factors data collection

Framework for development

- **Collect data on clinical practice**
- **Pre-emptive algorithm design OR compare existing algorithm to clinical practice**
- **Initial application**
- **Validation**
- **Refinement**

- **Study:**
- 3 hospital in 3 countries
- Achieved data extraction
- IT & clinical staff involved
- SSI after cardiac surgery, Colon surgery and hip/knee

UMC Utrecht

Example application of development framework

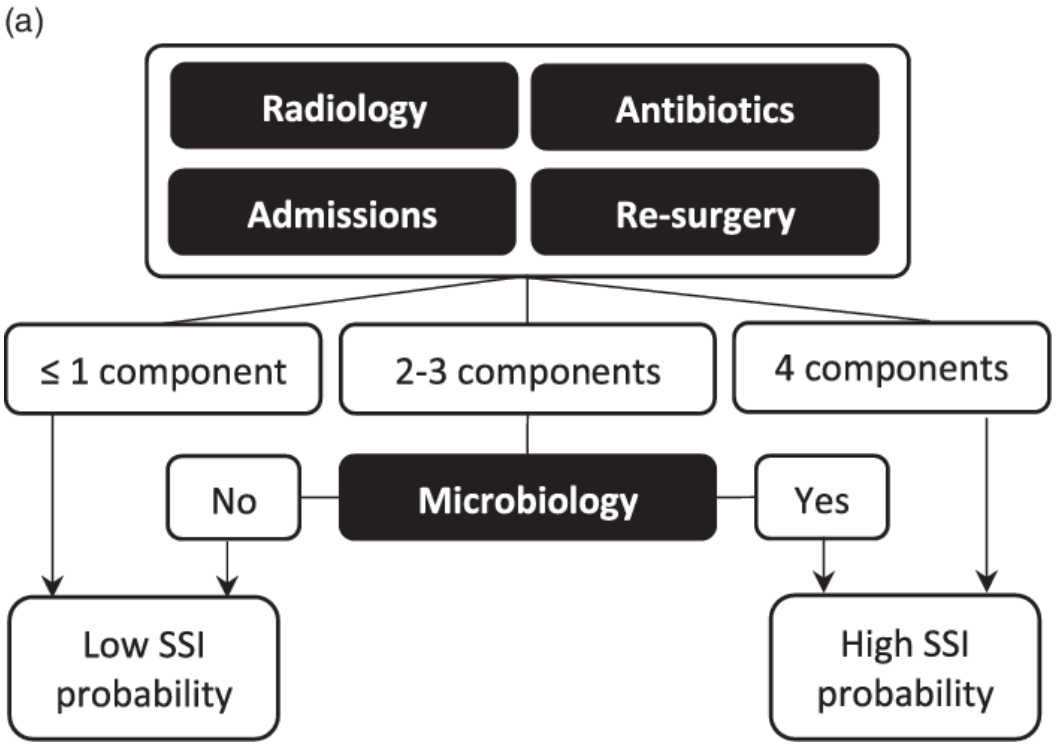
Do not over-specify an algorithm
Allow room for practice variation

Surgical Procedure	Antibiotics Included Algorithm	Hospital	Standardized Algorithm, % (No./Total)[a]			Center-Specific Algorithm, % (No./Total)[a]		
			Sensitivity[b]	PPV[c]	Workload Reduction[d]	Sensitivity[b]	PPV[c]	Workload Reduction[d]
Hip/knee prosthesis	Antibiotics	A	100.0 (8/8)	17.4 (8/47)	96.9 (47/1,509)	100.0 (8/8)	20.0 (8/40)	97.3 (40/1,509)
		B	83.3[e] (5/6)	62.5 (5/8)	97.5 (8/326)	50.0 (3/6)	37.5 (3/8)	97.5 (8/326)
	No antibiotics data	B	81.8[e] (9/11)	42.9 (9/21)	96.9 (21/686)	81.8[e] (9/11)	9.8 (9/92)	86.6 (92/686)
		C	94.7[e] (18/19)	18.4 (18/98)	96.2 (98/2,575)	94.7[e] (18/19)	15.1 (18/119)	96.2 (119/2,575)
Cardiac surgery	Antibiotics	A	97.0 (32/33)	34.8 (32/92)	96.1 (92/2,333)	93.9 (31/33)	43.7 (31/71)	97.0 (71/2,333)
		B	66.7 (6 /9)	19.4 (6/31)	93.0 (31/440)	44.4 (4/9)	33.3 (4/12)	97.3 (12/440)
	No antibiotics data	B	100.0 (15/15)	7.9 (15/191)	73.7 (191/725)	93.3 (14/15)	19.7 (14/71)	90.2 (71/725)
		C	95.7[e] (44/46)	8.3 (44/531)	73.2 (531/1,989)	89.1 (41/46)	21.5 (41/191)	90.4 (191/1,989)
Colon surgery	Antibiotics and radiology ordering included	A	93.3 (83/89)	36.1 (83/230)	82.2 (230/1,293)	86.5 (77/89)	45.3 (77/170)	86.9 (170/1,293)
		B	100.0 (16/16)	30.2 (16/53)	73.6 (53/201)	56.3 (9/16)	42.9 (9/21)	89.6 (21/201)
	Antibiotics and radiology ordering not included	B	83.7 (36/43)	33.6 (36/107)	72.3 (107/386)	48.8 (21/43)	43.8 (21/48)	87.6 (48/386)
		C	93.9 (92/98)	16.6 (92/554)	75.1 (554/2,227)	76.5 (75/98)	27.9 (75/269)	87.9 (269/2,227)

UMC Utrecht

Van Rooden et al A framework to develop semiautomated surveillance of surgical site infections: An international multicenter study; ICHE 2020

Example: Validation semi-automated surveillance SSI after colorectal surgery



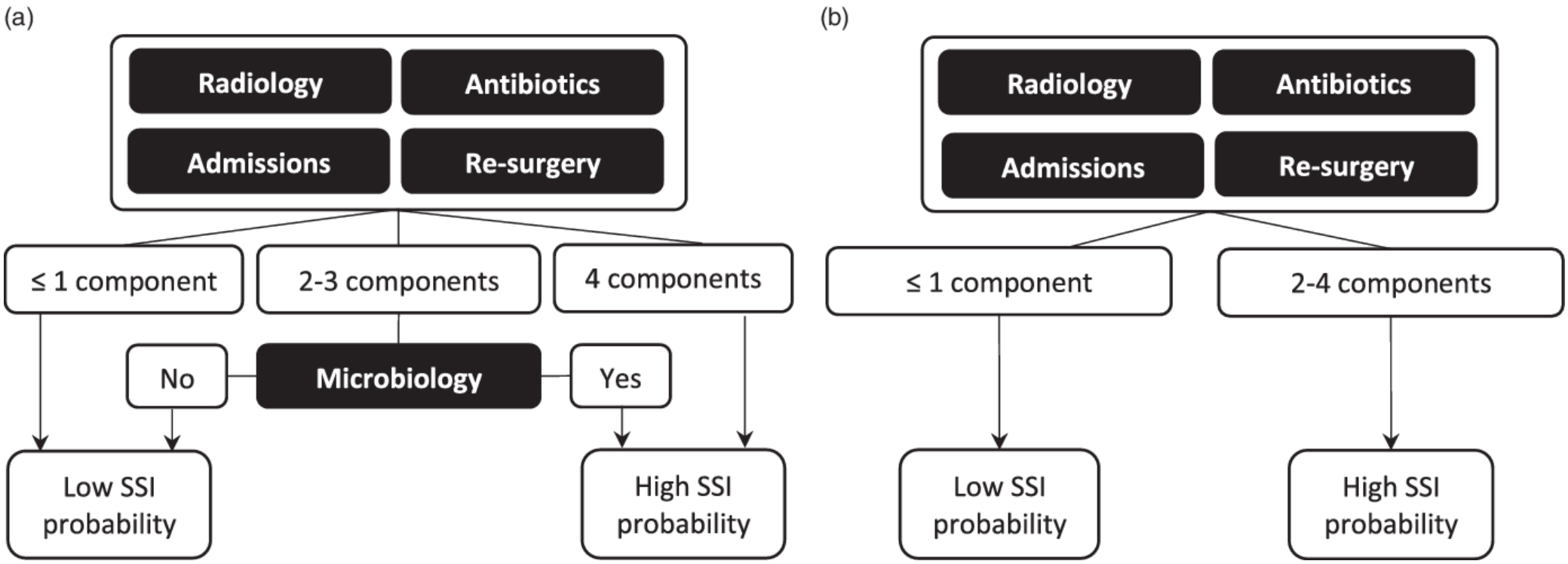
UMC Utrecht

Janneke Verberk et al, ICHE 2022

Validation prior to clinical alignment

Variable	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	% Reduction
Classification model					
Hospital A	100 (59.0–100.0)	90.4 (85.4–94.1)	26.9 (11.6–47.8)	100 (97.9–100.0)	87.4
Hospital B	100 (29.2–100.0)	89.3 (82.3–100.0)	18.8 (4.0–45.6)	100 (96.6–100.0)	87.2
Hospital C	85.7 (42.1–99.6)	92.2 (87.6–95.5)	27.3 (10.7–50.2)	99.5 (97.1–99.9)	89.7
Hospital D	72.7 (39.0–93.9)	97.5 (92.9–99.5)	72.7 (39.0–93.9)	97.5 (92.9–99.5)	91.6

Validation semi-automated surveillance SSI after colorectal surgery



UMC Utrecht

Janneke Verberk et al, ICHE 2022

After clinical alignment

Variable	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	% Reduction
Modified classification model					
Hospital A	100 (59.0–100.0)	77.8 (71.3–83.4)	13.7 (5.7–26.3)	100 (97.6–100.0)	75.2
Hospital B	100 (29.2–100.0)	80.1 (71.9–86.9)	11.1 (2.3–29.2)	100 (96.3–100.0)	78.3
Hospital C	100 (59.0–100.0)	77.6 (71.2–83.1)	13.2 (5.4–25.3)	100 (97.7–100.0)	75.0
Hospital D	100 (71.5–100.0)	89.2 (82.2–94.1)	45.8 (25.6–67.2)	100 (96.6–100.0)	81.7

Validate selection of surveillance population

Variable	Hospital A	Hospital B	Hospital C	Hospital D
Time period extractions	2019	2018–2019[a]	2019	2019[a]
Colorectal surgeries in reference standard, no.	205	167	221	142
Colorectal surgeries extracted automatically, no.	228	159	236	148
Matched records, no.	205	124	212	131
Deep SSI in matched records, no. (%)	7 (3.4)	3 (2.4)	7 (3.3)	11 (8.3)
Records in extractions that could not be linked to reference standard, no. (%)[b]	23 (10.1)	35 (22.0)	24 (10.2)	17 (11.4)
Records in reference standard that could not be linked to extractions, no. (%)[c]	0 (0.0)	43 (25.7)	9 (4.1)	11 (7.7)

b: Incorrect inclusion (non-primary)

c: Missed procedures: Operation by different specialty

Quality Quality Quality! Validate Validate Validate

UMC Utrecht

Steps in validation

Table 7
Validation requirements, at initiation and periodically, with examples

Characteristic	At initiation	Periodically (yearly)
Correct extraction of source data	Develop automated programming scripts to check for inconsistencies; outlier handling, technical validation.	Random sampling of data elements for manual verification.
	Manual verification of completeness by random sampling.	
Algorithm application	Assessment of completeness of coding systems (e.g. inclusion of relevant microbiologic results or antibiotics).	Monitor for changes in coding systems or IT updates.
	Programming errors.	
Algorithm performance	Assessment of algorithm to correctly identify patients with HAI (compare to reference standard).	Manual validation of a random or targeted sample.
	Agreement with clinical and documentation practices.	Audit of changes in clinical practice.
Denominator calculation	Correct application of inclusion and exclusion criteria (compared to references).	Manual validation.
	Calculation of device-days.	
Data sharing with (and analysis by) coordinating centre	Assessment integrity and completeness of data sent to coordinating centre.	Periodic manual check of data integrity and completeness.
Clinical acceptance	Discussion with clinicians.	Periodic discussion with clinicians.
	Association with other outcomes, if deemed relevant.	Associations with other outcomes.

Unless stated otherwise, these validation requirements apply to both locally and centrally implemented surveillance. Abbreviations: HAI, healthcare-associated infection; IT, information technology.

5. Shifting definitions

Many case definitions include unstandardised clinical information

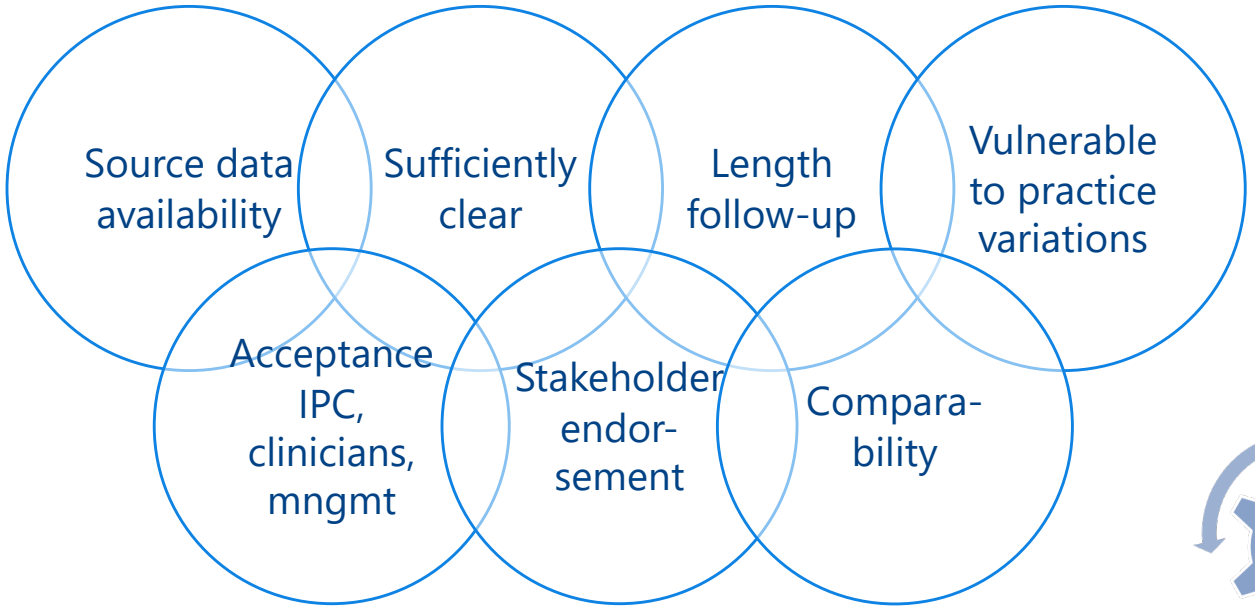
- Signs & symptoms
- Aspect of wounds, abcesses
- Radiological description

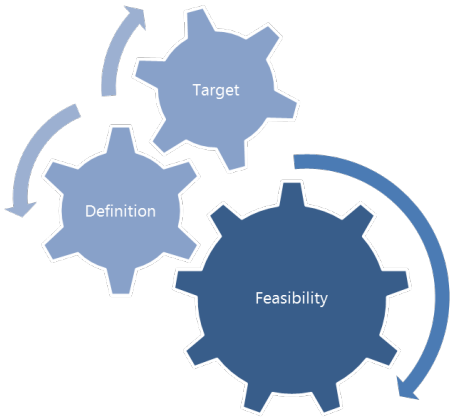
- **Semi-automated surveillance:**
- Manual ascertainment can correct (some) of this
- Sensitivity is key

- **Fully-automated surveillance**
- Must adapt definition

Design of AS (2)

✓Automated surveillance requires reconsideration of HAI case definitions to address limitations in data availability and methodological aspects of case-ascertainment





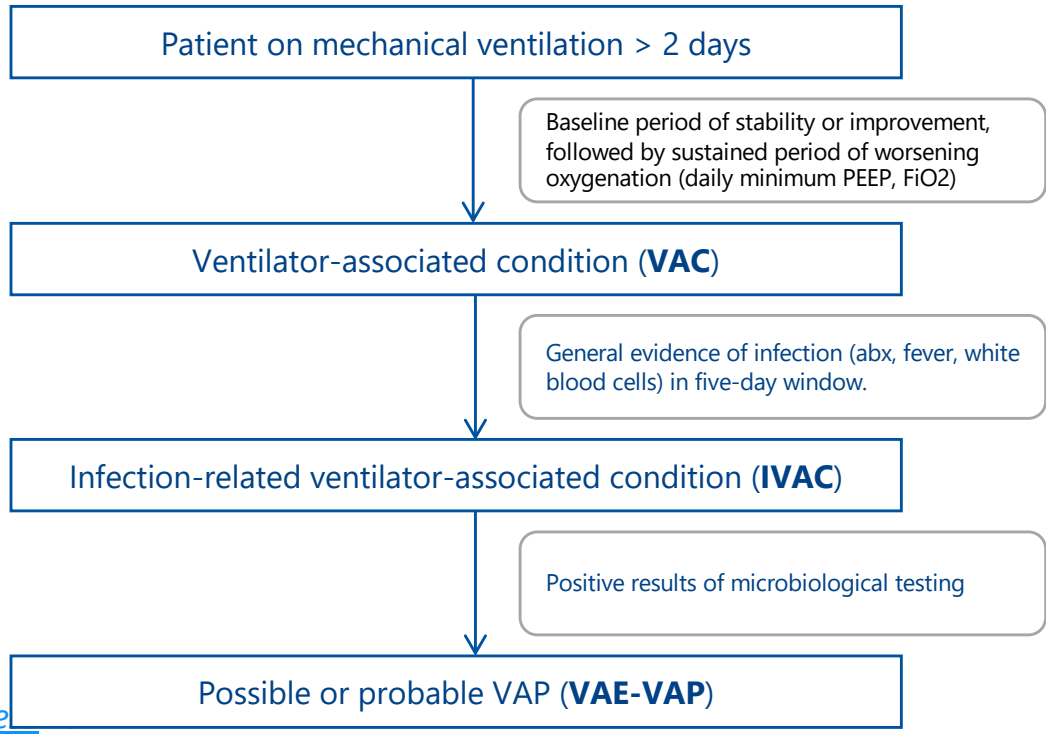
Shifting definitions: Ventilator-associated events

Remove subjectivity and facilitate automated implementation

Ventilator settings, no 'human interpretation'

Use of electronic data does not guarantee comparability.
Vulnerability to manipulation remains

Changing entities complicates interpretation
Broad scope of conditions: ARDS, fluid overload, pneumonia, …
Preventable events?
What is effect of case-mix
What actions to take if the rate is high?



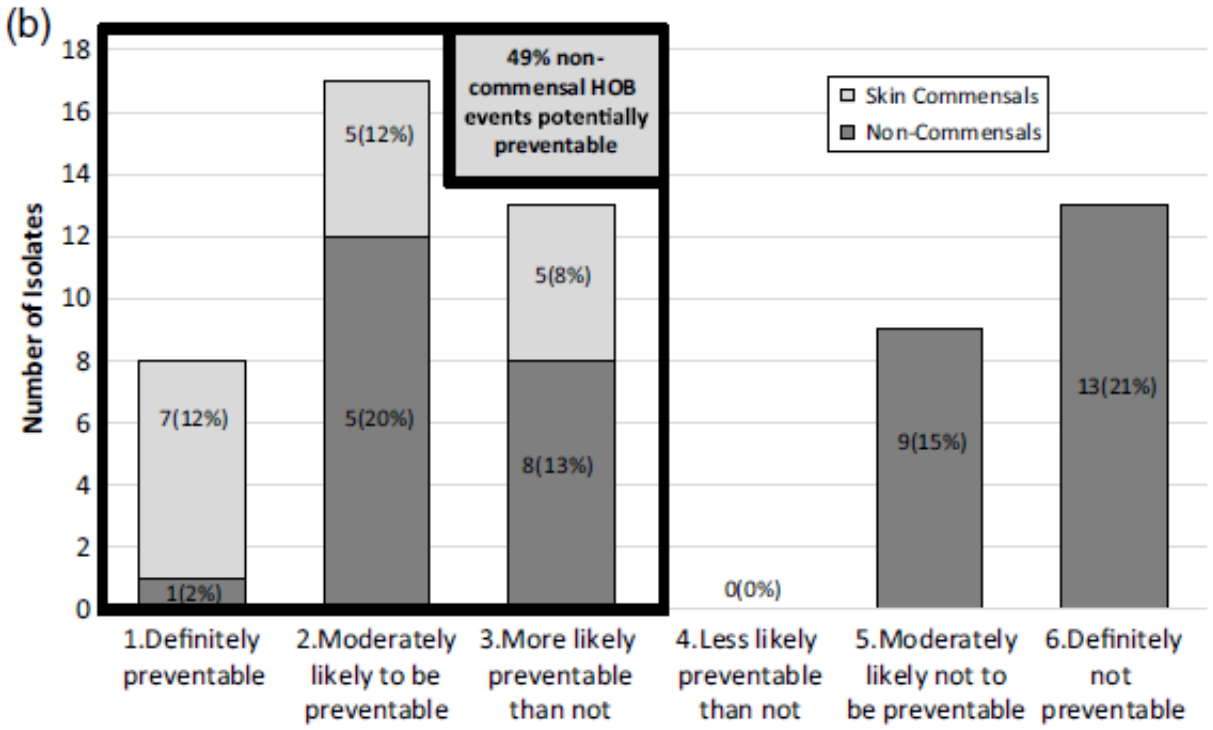
Klein Klouwenberg *Am J Resp Crit Care* 2014, Lilly et al 2014, Magill et al. *Curr Opin Infect Dis* 2014, Boyer *Che*

Example: Hospital-onset bacteremia

DISCLAIMER – UNDER DEVELOPMENT

U.S.

- Any positive bloodculture > 48 hours after admission
- Correlation with CLABSI rate (1 per 1000 PD increase in HOB -> 2,5% relative increase in CLABSI)
- Overlap with CLABSI: 6-20%
- Common skin commensals: 13%



Judged partially preventable

No studies assessing interventions

PRAISE Network: Definition under development

Rock *ICHE* 2016, Dantes *ICHE* 2019, PRAISE Network

TABLE 1. ICU Types, Frequencies, and Rates of Central-Line–Associated Bloodstream Infection (CLABSI) and Hospital–Onset Bacteremia (HOB)

ICU Type	No. ICU	Total No. CLABSI	Total Central-Line Days	CLABSI Rate[a]	No. CLABSIs, Range	CLABSI Rate, Range[a]	Total No. HOB	Total No. ICU Patient Days	HOB Rate[b]	No. HOB, Range	HOB Rate, Range[b]
Medical	12	104	85,858	1.21	1–19	0.29–3	2,735	152,404	17.95	73–402	9.41–39.89
Cardiac	10	53	43,234	1.23	1–13	0.21–3.77	1,254	78,869	15.90	35–216	3.54–38
Surgical	10	77	69,100	1.11	2–23	0.19–2.36	1,621	127,936	12.67	46–251	5.42–24.84
Neonatal	9	99	76,139	1.30	2–15	0.45–2.33	776	238,921	3.25	37–156	1.12–9.27
Pediatric: Medical/ Surgical	9	78	40,300	1.94	0–20	0–4	880	88,601	9.93	7–203	2.59–18.3
Cardiothoracic	7	64	57,919	1.10	0–17	0–1.7	972	76,604	12.69	14–327	4.07–28.67
Trauma	6	57	28,867	1.97	2–17	0.8–2.68	888	56,133	15.82	120–171	8.25–22.05
Neurosurgical	5	29	26,369	1.10	1–11	0.14–2.57	460	66,469	6.92	65–136	4.77–10.1
Burn	4	38	7,426	5.12	1–24	0.86–11.23	346	24,454	14.15	38–145	6.88–40.41
Medical/Surgical	4	35	19,471	1.80	0–23	0–2	710	32,082	22.13	17–414	7.65–27.16
Neurologic	2	4	7,864	0.51	0–4	0–0.74	269	22,037	12.21	119–150	9.51–18.96
Pediatric: Cardiothoracic	1	13	7,266	1.79	13–13	1.79–1.79	87	8,162	10.66	87–87	10.67–10.67
Pediatric: Mixed Acuity Unit	1	12	5,607	2.14	12–12	2.14–2.14	282	9,934	28.39	282–282	28.39–28.39
Total for all ICUs	80	663	475,420				11,280	982,609			

71.7% of ICU-months with zero events

11.5% of ICU-months with zero events

Food for thought!

UMC Utrecht

6. Risks of automated surveillance

Change in methodology is not without consequences

- Changing definitions -> changing interpretation & break in data
- AS data ≠ manually collected data
- Risk of losing comparability amongst networks if different methods are chosen.

Assessment of value of AS in delivering data for quality improvement

AS is not a guarantee for comparability

- Data sources, underlying clinical practice, technical implementation
- Maintenance

Concluding remarks & THM

Automated surveillance has potential to improve quality & efficiency of surveillance

Requires accessible source (EHR) data of sufficient quality and consistency

Development of algorithms requires
- Clinical validation(s)
- Sometimes modification of definitions

Many approaches to implementation, also depending on purpose
- Fully vs. Semi-automated
- Central vs. Local implementation

Questions?

M.S.M.vanMourik-2@umcutrecht.nl

www.webbertraining.com/schedulep1.php

March 9, 2023	HOMECARE & HOSPICE - STANDARDIZING INFECTION SURVEILLANCE Speaker: **Mohamed Adawee**, Sparrow Health, Michigan
March 23, 2023	THE ENVIRONMENT, THE TICK, AND THE PATHOGEN - IT'S AN ENSEMBLE Speaker: **Jannelle Couret**, University of Rhode Island
April 4, 2023	*(FREE European Teleclass)* RESPIRATORY INFECTION PREVENTION: PERCEPTIONS, BARRIERS AND FACILITATORS Speaker: **Dr. Pierre Parneix**, Hôpital Pellerin, CHU de Bordeaux, France
April 12, 2023	*(South Pacific Teleclass)* UNINTENDED CONCEQUENCES OF INFECTION PREVENTION AND CONTROL MEASURES DURING THE COVID-19 PANDEMIC Speaker: **Dr. Moi-Lin Ling**, SingHealth, Singapore
April 20, 2023	HOSPITAL WASTEWATER SYSTEMS: ORIGINS OF NOVEL NOSOCOMIAL BACTERIA Speaker: **Professor Colum Dunne**, School of Medicine, University of Limerick, Ireland
April 27, 2023	THE FUNGUS AMONG US: THE EMERGENCE OF A HIGHLY RESISTANT FUNGUS IN THE HEALTHCARE SYSTEM Speaker: **Dr. Tom Chiller**, Centers for Disease Control, Atlanta
May 5, 2023	*(FREE Teleclass)* SPECIAL LECTURE FOR 5 MAY Speaker: **Prof. Didier Pittet**, University of Geneva, Switzerland
	(European Teleclass)

Thanks to Teleclass Education

PATRON SPONSORS

diversey.com

virox.com

gamahealthcare.com