Artificial Intelligence in Medicine 69 (2016) 33–41

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Detecting borderline infection in an automated monitoring system for healthcare-associated infection using fuzzy logic

Jeroen S. de Bruin[a,*], Klaus-Peter Adlassnig[a,b], Alexander Blacky[c], Walter Koller[d]

[a] Section for Artificial Intelligence and Decision Support, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria
[b] Medexter Healthcare GmbH, Borschkegasse 7/5, A-1090 Vienna, Austria
[c] VAMED-KMB Hospital Management and Operation GmbH, Spitalgasse 23, A-1090 Vienna, Austria
[d] University Clinic for Hospital Hygiene and Infection Control, Medical University of Vienna and Vienna General Hospital, Waehringer Guertel 18-20, A-1090 Vienna, Austria

ARTICLE INFO

Article history:
Received 3 February 2016
Received in revised form 27 April 2016
Accepted 27 April 2016

Keywords:
Automated monitoring and surveillance systems
Infection control
Fuzzy logic
Intensive care units
Healthcare-associated infections

ABSTRACT

Background: Many electronic infection detection systems employ dichotomous classification methods, classifying patient data as pathological or normal with respect to one or several types of infection. An electronic monitoring and surveillance system for healthcare-associated infections (HAIs) known as Moni-ICU is being operated at the intensive care units (ICUs) of the Vienna General Hospital (VGH) in Austria. Instead of classifying patient data as pathological or normal, Moni-ICU introduces a third *borderline* class. Patient data classified as borderline with respect to an infection-related clinical concept or HAI surveillance definition signify that the data nearly or partly fulfill the definition for the respective concept or HAI, and are therefore neither fully pathological nor fully normal.
Objective: Using fuzzy sets and propositional fuzzy rules, we calculated how frequently patient data are classified as normal, borderline, or pathological with respect to infection-related clinical concepts and HAI definitions. In dichotomous classification methods, borderline classification results would be confounded by normal. Therefore, we also assessed whether the constructed fuzzy sets and rules employed by Moni-ICU classified patient data too often or too infrequently as borderline instead of normal.
Participants and methods: Electronic surveillance data were collected from adult patients (aged 18 years or older) at ten ICUs of the VGH. All adult patients admitted to these ICUs over a two-year period were reviewed. In all 5099 patient stays (4120 patients) comprising 49,394 patient days were evaluated. For classification, a part of Moni-ICU's knowledge base comprising fuzzy sets and rules for ten infection-related clinical concepts and four top-level HAI definitions was employed. Fuzzy sets were used for the classification of concepts directly related to patient data; fuzzy rules were employed for the classification of more abstract clinical concepts, and for top-level HAI surveillance definitions. Data for each clinical concept and HAI definition were classified as either normal, borderline, or pathological. For the assessment of fuzzy sets and rules, we compared how often a borderline value for a fuzzy set or rule would result in a borderline value versus a normal value for its associated HAI definition(s). The statistical significance of these comparisons was expressed in p-values calculated with Fisher's exact test.
Results: The results showed that, for clinical concepts represented by fuzzy sets, 1–17% of the data were classified as borderline. The number was substantially higher (20–81%) for fuzzy rules representing more abstract clinical concepts. A small body of data were found to be in the borderline range for the four top-level HAI definitions (0.02–2.35%). Seven of ten fuzzy sets and rules were associated significantly more often with borderline values than with normal values for their respective HAI definition(s) ($p < 0.001$).
Conclusion: The study showed that Moni-ICU was effective in classifying patient data as borderline for infection-related concepts and top-level HAI surveillance definitions.

* Corresponding author at: Section for Artificial Intelligence and Decision Support, Center for Medical Statistics, Informatics and Intelligent Systems (CeMSIIS), Medical University of Vienna, Spitalgasse 23, BT 88.03.714, A-1090 Vienna, Austria. Fax: +43 1 40400/66250.
E-mail address: jeroen.debruin@meduniwien.ac.at (J.S. de Bruin).

http://dx.doi.org/10.1016/j.artmed.2016.04.005
0933-3657/

1. Introduction

Electronic systems for the detection and monitoring of healthcare-associated infections (HAIs) have become common in clinical routine over the last decade [1–3]. Electronic monitoring is considered superior to traditional surveillance because electronic systems are faster, require less human resources, and are not subject to inter-rater variability as manual surveillance is [3–6].

A common limitation of electronic infection monitoring systems is that they employ dichotomous "yes/no" classification methods, thereby classifying an infection as either present or absent. Consequently, *borderline* infection cases, i.e., patients who show vital signs and laboratory test results that nearly or partially fulfill the conditions defined in the infection surveillance rules are not explicitly recognized. Instead, they are confounded by patients whose vitals and test results are normal. Borderline infection cases are clinically relevant, as these patients are at a high risk of developing infection, need close monitoring and possibly interventions as well. As such, failing to identify borderline infection cases reduces the usefulness of electronic infection monitoring systems in predicting, alerting, and preventing infection.

One way of differentiating borderline infection cases from patients without signs of infection is through the integrated use of fuzzy sets and logic [7]. By using fuzzy sets for the formal representation of infection-related clinical concepts such as fever, hypertension, or leukocytosis, we extend the traditional dichotomous classification methods. An accurately defined fuzzy set can classify patient data as *not compatible* (normal), *fully compatible* (pathological), or expressing a *degree of compatibility* (nearly or partly pathological, or borderline) with respect to a clinical concept under consideration. In the latter case, a degree of compatibility between measured or observed patient data and the respective clinical concept represents a gradual transition from normal to pathological values for the respective concept. After the initial evaluation of clinical concepts using fuzzy sets, fuzzy logic is used to evaluate logical combinations of these concepts in order to draw conclusions about higher-level concepts, and ultimately infer the full or partial compatibility or non-compatibility of the top-level HAI terms with the underlying patient data.

Electronic HAI monitoring systems using fuzzy sets and logic remain as effective as their non-fuzzy counterparts in the detection of pathological infection cases [8]. However, they possess the additional ability to permit a distinction between patients with a suspected borderline infection and normal patients. This ability offers the following advantages: (a) more accurate feedback on the patients' status to the attending physicians and the infection control experts, and (b) identifying incipient and recurring infections [9], which allows early therapeutic intervention.

The goal of the present study was to separate borderline infection cases from patients without signs of infection, and assess the size of the newly created patient group with respect to the aforementioned two patient classes (normal and pathological). Using the electronic HAI surveillance and monitoring system Moni-ICU [10–12], we calculated the frequencies of patient data in the categories of normal, borderline, or pathological for all of the incorporated fuzzy sets. The results of the application of fuzzy rules for higher-level clinical concepts and HAI surveillance definitions were calculated in the same manner. We also analyzed the present definitions of fuzzy sets and rules to ensure they did not classify patient data too often or too infrequently as borderline rather than normal. To this end, we postulated the hypothesis that borderline values for a clinical sign or symptom should more often result in a borderline value than in a normal value for its associated HAI definition(s). By constructing contingency tables for each fuzzy set and fuzzy rule and their associated HAI definition(s), we were able to express the ability of a fuzzy set or fuzzy rule in separating borderline infection cases from normal patients with the aid of p-values.

2. Background

2.1. Fuzzy set theory and fuzzy logic

Fuzzy set theory and fuzzy logic are being developed since 1965. Fuzzy sets have been introduced to express partial membership of objects to classes, which are usually characterized by their linguistic terms. A so-called *degree of membership* [13] indicates the degree to which the linguistic term is present; it expresses the degree of compatibility between the measured underlying value and the concept under consideration. Subjectively interpretable linguistic clinical concepts are commonly used in medical definitions, protocols, and guidelines. Fuzzy sets can be employed to model the unsharpness of clinical terms when trying to diagnose a patient's condition on the basis of his/her medical data [14,15]. This process includes the calculation of compatibility between measured patient data and the linguistic clinical concept under consideration (Fig. 1).

Fuzzy logic can be used to evaluate logical combinations of concepts that are assigned a degree of membership [16]. In the present study, we employed propositional fuzzy logic, a many-valued logic used to reason and make inferences about one or more evaluated fuzzy sets (including the results of the evaluation of crisp sets as a specialization of fuzzy sets). We employed three propositional fuzzy operators throughout this report: conjunction, disjunction, and negation.

Conjunction is commonly interpreted by a t-norm $\odot$: $[0,1]^2 \rightarrow [0,1]$. Any t-norm is associative, commutative, neutral with respect to 1, and isotone in both arguments [17]. Disjunction is usually interpreted by the corresponding t-conorm $\oplus$: $[0,1]^2 \rightarrow [0,1]$. For the present study we used the Gödel t-norm $\odot_G$, the associated Gödel t-conorm $\oplus_G$, and the standard negation function $\neg$: $[0,1] \rightarrow [0,1]$. Given x, y $\in$ [0,1], these are defined as follows:

$$x \odot_G y = \min(x, y)$$

$$x \oplus_G y = \max(x, y)$$

$$\neg x = 1 - x$$

Fuzzy set theory and fuzzy logic have become increasingly popular in medicine over the last thirty years, especially in the areas of fuzzy classification and inference [18,19]. Abbod et al. [20] provide an overview of applications using fuzzy sets and/or fuzzy logic in many specialties disciplines of medicine while Mahfouf et al. [21] present a comprehensive survey on applications in medicine that use fuzzy logic for monitoring and control.

2.2. Healthcare-associated infections

An ICU-based HAI is defined as an infection manifested in a patient later than 48 h after admission to the ICU. Electronic HAI monitoring is based on the ICU surveillance rules defined by the European Centre for Disease Prevention and Control (ECDC) surveillance program. Several definitions of infection are included therein [22]:

- Blood stream infection (BSI),
- Pneumonia (PN1-5),
- Central venous catheter-related infection (CRI1-2),
- Urinary tract infection (UTI-A and -B).

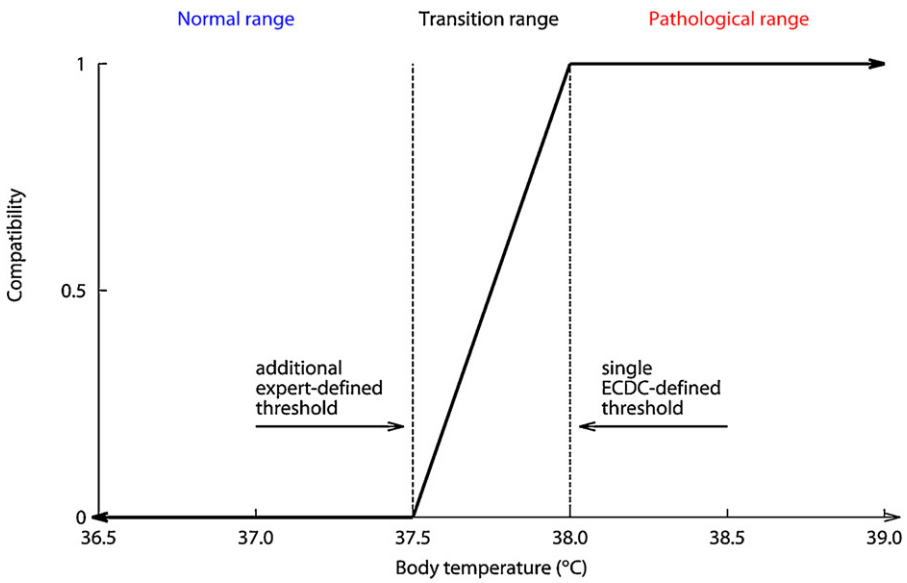

Fig. 1. Graphic depiction of a fuzzy set for the linguistic clinical concept *increased body temperature*. The *x*-axis shows the range of human body temperature (°C) while the *y*-axis shows the compatibility of a specific patient's body temperature with the concept *increased body temperature*. The range of input values that yields a degree of compatibility between zero and one (transition range) is between 37.5 °C and 38 °C. 38 °C is the threshold for fever as defined by the European Centre for Disease Prevention and Control (ECDC) for its HAI surveillance program. The 37.5 °C threshold was established by clinical experts for the purpose of early intervention.

In the present study, we analyzed patient data in monitoring BSI, CRI2, UTI-A, and UTI-B. Pneumonia was omitted because radiology findings are not yet available for automated import into Moni-ICU at the Vienna General Hospital (VGH). CRI1 was also not included because its present definition in the knowledge base of Moni-ICU does not involve fuzzy sets or fuzzy rules.

3. Methods

3.1. Study setting and design

Data from ten ICUs at the VGH, a 2133-bed tertiary care and teaching hospital run jointly with the Medical University of Vienna (MUV) were used for the study, which was reviewed and approved by the ethics committee of MUV (approval number 1888/2012).

3.2. Participants and study period

All adult patients (aged 18 or older) present in, or admitted to, one of the ten selected ICUs between 1 January 2011 and 31 December 2012 were eligible for the study. In accordance with the ECDC surveillance guidelines on HAIs, patients staying for 48 h or less in an ICU were excluded.

3.3. Data sources

Moni-ICU imports clinical, laboratory, and nursing data from the Philips IntelliSpace Critical Care and Anesthesia (ICCA) patient data management system, which is employed in the ICUs of the VGH. For microbiology data, Moni-ICU is connected with the MOLIS laboratory information system at the Clinical Department of Clinical Microbiology. Data from the two sources are consolidated by unique patient and ward admission identifiers.

3.4. Knowledge base and data processing

Moni-ICU, an electronic HAI monitoring system developed jointly by VGH and MUV, was used for data processing. Moni-ICU is being used since 2004 as a support tool for infection monitoring and surveillance as well as research projects.

The knowledge base consists of computerized, formal representations of ECDC's HAI surveillance definitions implemented in Arden Syntax 2.7. Arden Syntax is a programming language for encoding medical knowledge in both machine- and human-readable form. The sequence and conditions of processing input and output data are specified in so-called medical logic modules (MLMs) [23,24]. Each MLM in the knowledge base is executed once a day for each patient. If necessary, the defined MLMs can take input data accumulated over a period of seven days into account.

3.4.1. Fuzzy sets

Fuzzy sets are used to represent infection-related clinical concepts for which compatibility can be determined directly from the available measured raw data. Fuzzy sets for six clinical concepts are taken into account: *increased body temperature*, *shock*, *drop in blood pressure*, *increased C-reactive protein*, *leukopenia*, and *leukocytosis* (Table 1).

The aforementioned fuzzy sets were defined by infection control experts together with clinical knowledge engineers. For each fuzzy set, numerical thresholds for pathological values as defined by the ECDC surveillance rules were used. When a threshold was not available, it was determined by clinical expertise. A second threshold signifying the end of the range for normal values was also defined. The space between the two thresholds is referred to as the transition range, in which degrees of compatibility are determined by a monotonically increasing linear function.

3.4.2. Fuzzy rules

Several concepts in Moni-ICU represent aggregated, higher-level clinical terms that cannot be directly derived from raw measured data (Table 2). These concepts include *fever incorporating thermoregulation*, *hypotension/shock*, *clinical signs of BSI*, and *clinical signs of UTI*.

The four top-level terms in ECDC's HAI surveillance definitions discussed in the present study are *BSI*, *CRI2*, *UTI-A*, and *UTI-B*.

Table 1
Overview of infection-related clinical concepts in Moni-ICU, represented by fuzzy sets.

Clinical concept (unit)	Fuzzy set		
	Normal range	Transition range	Pathological range
Increased body temperature (°C)	[0,37.5]	]37.5, 38[[a]	[38,+∞] [b]
Shock (systolic blood pressure/heart rate)	[1.3, +∞]	]1, 1.3[[a]	[0,1] [b]
Drop in blood pressure (lower percentile difference of average over two days)	[100,37.5]	]37.5, 25[[a]	[25,0] [b]
Increased C-reactive protein (mg/dl)	[0,1]	]1, 6[[c]	[6,+∞] [b]
Leukopenia (WBC/mm^3)	[5000,+∞]	]4,000, 5,000[[d]	[0,4,000] [b]
Leukocytosis (WBC/mm^3)	[0,11000]	]11,000, 12,000[[d]	[12,000,+∞] [b]

Note: [] indicates thresholds are included in the interval;] [indicates thresholds are excluded from the interval; WBC, white blood cell.
[a] As defined by clinicians.
[b] As defined by the CDC/NHSN [25], ECDC [22], and KISS [26] infection surveillance programs for retrospective surveillance purposes.
[c] As defined by clinicians; C-reactive protein is an early-phase protein, useful for prospective detection purposes.
[d] As defined by clinicians; WBC is a slow-reacting indicator, important for surveillance purposes.

Table 2
Overview of infection-related clinical concepts in Moni-ICU, represented by fuzzy rules.

Clinical concept	Fuzzy rule or HAI definition
Fever incorporating thermoregulation	*Increased body temperature* ⊕ (0.8 * *#thermoregulation*)
Hypotension/shock	0.8 * (*Drop in blood pressure* ⊕ *shock*)
Clinical signs of BSI	*Fever incorporating thermoregulation* ⊕ *hypotension/shock* ⊕ *increased C-reactive protein* ⊕ *leukopenia* ⊕ *leukocytosis*
Clinical signs of UTI	*Fever incorporating thermoregulation* ⊕ *increased C-reactive protein* ⊕ *leukopenia* ⊕ *leukocytosis*
BSI	*#PBC* ⊕ (*#CSC* ⊙ *Clinical signs of BSI*)
CRI2	(¬*#PBC*) ⊙ *Clinical signs of BSI* ⊙ *#PCC*
UTI-A	*Clinical signs of UTI* ⊙ *#PUC*
UTI-B	*Clinical signs of UTI* ⊙ *#diagnostic clue(s) for UTI other than PUC*

Note: # indicates crisp concepts; ⊙ fuzzy conjunction; ⊕ fuzzy disjunction; ¬ fuzzy negation; HAI, healthcare-associated infection; *PBC*, positive blood culture; *CSC*, two or more separate blood cultures positive for the same common skin contaminant; *PCC*, positive catheter culture; *PUC*, positive urine culture; *BSI*, blood stream infection; *UTI*, urinary tract infection; *CRI*, central venous catheter-related infection.

Compatibility between patient data and the above concepts is determined by rules that apply fuzzy logic operators on previously evaluated concepts and rules (both crisp and fuzzy).

For some rules, compatibility is adjusted by a numerical modifier (such as multiplication by 0.8). This modifier reflects uncertainty in data or medical knowledge, or the influence of interventions. As an example, consider the fuzzy rule for the concept *fever incorporating thermoregulation* (cf., Table 2). It partly depends on the underlying concept *increased body temperature,* but also on the crisp concept *thermoregulated*. Patients with exceptionally high fever are cooled with cooling pads or blankets. While this intervention implies that the patient had high fever at one point in time, the patient's elevated body temperature cannot be measured as long as thermoregulation is in progress. Therefore, any conclusion will be indirect, and the compatibility for *thermoregulated* is adjusted using the aforementioned modifier.

To illustrate how the Moni-ICU system functions, consider the patient data of *Patient X*, shown in Table 3. First, the values for clinical concepts represented by fuzzy sets are determined (cf., Table 1). In this example, borderline values are recorded for *drop in blood pressure* (0.48), *increased body temperature* (0.6), and *increased C-reactive protein* (0.56); all other input values are in the normal range and therefore yield 0.

After all fuzzy sets have been processed, values for fuzzy rules representing higher-level clinical concepts (cf., Table 2) are calculated using standard negation, the Gödel t-norm and t-conorm. Values for *hypotension/shock* and *fever incorporating thermoregulation* are calculated first. For *hypotension/shock*, the maximum value for either *shock* (0) or *drop in blood pressure* (0.48) is chosen, and multiplied by 0.8. The resulting value is thus $0.48 \times 0.8 = 0.384$. For *fever incorporating thermoregulation*, the highest value is chosen between *increased body temperature* (0.6) and (*#thermoregulation* * 0.8). As thermoregulation was not applied to *Patient X*, the value is 0.6. Next, the values for *clinical signs of BSI* and *clinical signs of UTI* are determined. Both rules propagate the maximum value of several parameters previously calculated, including *fever incorporating thermoregulation*. As this parameter has the highest calculated value (0.6), it is propagated to both parameters.

After calculation of higher-level clinical concepts, values for the top-level HAI definitions are determined. As no common skin contaminants nor any positive blood or urine cultures were present, the value for *BSI* and *UTI-A* equals zero (cf., Table 2). Furthermore, as no other diagnostic clues for UTI were present, *UTI-B* is also ruled out. However, since there was no positive blood culture, a non-zero value for the parameter *clinical signs of UTI*, and a positive catheter culture, the value for the *CRI2* rule (cf. Table 2) equals min ((1 − 0), 0.6, 1)) = 0.6. Thus we conclude that *Patient X* shows borderline signs for *CRI2*. A workflow for this example is shown in Fig. 2.

3.5. Outcome measures

For each fuzzy set or rule, we present classification frequencies (patient days) for each class (normal, borderline, pathological). Since borderline patient data would be confounded by normal patient data in dichotomous classification methods, we also specified a reclassification ratio (RR). This ratio expresses the relative frequency of borderline patient data with reference to the frequency of normal patient data before reclassification, defined as:

$$\frac{\text{patient days with borderline patient data}}{\text{patient days with borderline patient data} + \text{patient days with normal patient data}}$$

whereby RR ∈ [0,1]. This ratio can be interpreted as follows: if RR yields zero it means there were no borderline patient data; if RR yields one it means all patient data were borderline; values between zero and one indicate the fraction of data previously clas-

Table 3
Example patient data for *Patient X*.

Clinical data		Microbiology data	
Body temperature	37.8 °C	Common skin contaminants	no
Shock index	1.4	Positive blood culture	no
Blood pressure index	31	Positive catheter culture	yes
C-reactive protein level	3.8 mg/dl	Positive urine culture (PUC)	no
White blood cell count	7,000/mm^3		
Thermoregulation	No		
Diagnostic clues for UTI other than PUC	No		

Note: UTI, urinary tract infection.

Table 4
Frequency distributions for infection-related clinical concepts, modeled by fuzzy sets or fuzzy rules.

Clinical concept	Frequency (%)			RR
	Normal	Borderline	Pathological	
Fuzzy sets				
Increased body temperature	35,839 (72.56)	3,235 (6.55)	10,320 (20.89)	0.083
Shock	20,270 (41.04)	8,056 (16.31)	21,068 (42.65)	0.284
Drop in blood pressure	20,849 (42.21)	3,805 (7.70)	24,740 (50.09)	0.154
Increased C-reactive protein	13,710 (27.76)	5,939 (12.02)	29,745 (60.22)	0.302
Leukopenia	47,214 (95.59)	642 (1.30)	1,538 (3.11)	0.013
Leukocytosis	30,792 (62.34)	1,666 (3.37)	16,936 (34.29)	0.051
Fuzzy rules				
Fever incorporating thermoregulation	7,481 (15.15)	31,593 (63.96)	10,320 (20.89)	0.809
Hypotension/shock	9,742 (19.72)	39,652 (80.28)	–	0.803
Clinical signs of BSI	218 (0.44)	12,019 (24.33)	37,157 (75.23)	0.982
Clinical signs of UTI	2,234 (4.52)	10,003 (20.25)	37,157 (75.23)	0.817
HAI surveillance definitions				
BSI	47,787 (96.74)	8 (0.02)	1,599 (3.24)	<0.001
CRI2	43,633 (88.34)	1,161 (2.35)	4,600 (9.31)	0.026
UTI-A	48,559 (98.31)	232 (0.47)	603 (1.22)	0.005
UTI-B	49,330 (99.87)	13 (0.03)	51 (0.10)	<0.001

Note: RR, reclassification ratio; BSI, blood stream infection; CRI, Central venous catheter-related infection; UTI, urinary tract infection.

sified as normal, that were re-classified as borderline using fuzzy sets and rules.

For the assessment of fuzzy sets and rules, data are presented in 2×2 contingency tables with the variables "clinical concept class" (borderline/normal) and "top-level HAI rule class" (borderline/normal). The upper-left field indicates all patient days for which values for both the clinical infection-related concept and the associated top-level HAI rule(s) were borderline. The upper-right field indicates all patient days for which the value for the clinical infection-related concept was borderline, but the value for the associated top-level HAI rule(s) was normal. The lower-left field indicates all patient days for which the value for the clinical infection-related concept was normal, but the value for the associated top-level HAI rule(s) was borderline. Finally, the lower-right field indicates all patient days for which values for both the clinical infection-related concept and the associated top-level HAI rule(s) were normal. Based on these contingency tables, we used Fisher's exact test for comparisons. The level of significance was set to $p < 0.05$.

3.6. Data analysis

Data filtering and cleaning were performed with Microsoft Excel 2013. Study outcome measures and p-values were calculated in R using the function for Fisher's exact test for count data.

4. Results

In the study period, 4120 patients stayed at one or more of the selected ICUs; 2454 patients were male (61.04%), and 1654 female (41.14%); gender was not recorded for twelve patients. The patients' ages ranged between 18–98 years, with a median of 61 years and an interquartile range of 23. In all 5099 patient stays were recorded, comprising 49,394 patient days. The duration of the hospital stay ranged between 3–187 days, with a median of six days and an interquartile range of ten.

On average 56.91% (27.76–95.59) of the data related to clinical concepts modeled by fuzzy sets were classified as normal, 7.88% (1.30–16.31) as borderline, and 35.21% (3.11–60.22) as pathological (Table 4). Depending on the clinical concept, RR yields 0.01–0.30, meaning that 1–30% of the data previously classified as normal were reclassified into the borderline class.

For fuzzy rules, there were markedly more borderline classifications and significantly fewer normal ones. On average 9.95% (0.44–19.72) of the data were rated normal, 47.21% (20.25–80.28) borderline, and 42.84% (0–75.23) pathological. For the higher-level clinical symptoms of *fever incorporating thermoregulation* and *hypotension/shock*, most data were classified as borderline (63.96–80.28%). For rules aggregating clinical signs, most data were classified as pathological (75.23%). Reclassification ratios were high; 80-98% of the data were reclassified into the borderline class.

Finally, the HAI surveillance definitions revealed very few borderline or pathological results. On average 97.61% (88.34–99.87) of the data were classified as normal, whereas only 0.41% (0.02–2.35) were classified as borderline and 1.98% (0.10–9.31) as pathological. The highest number of borderline and pathological classifications were registered for *CRI2*. The smallest number of borderline infection cases was noted for *BSI*, and the smallest number of pathological infections for *UTI-B*. As there were very few borderline infection cases and most data were in the normal range, reclassification ratios resembled borderline classification frequencies.

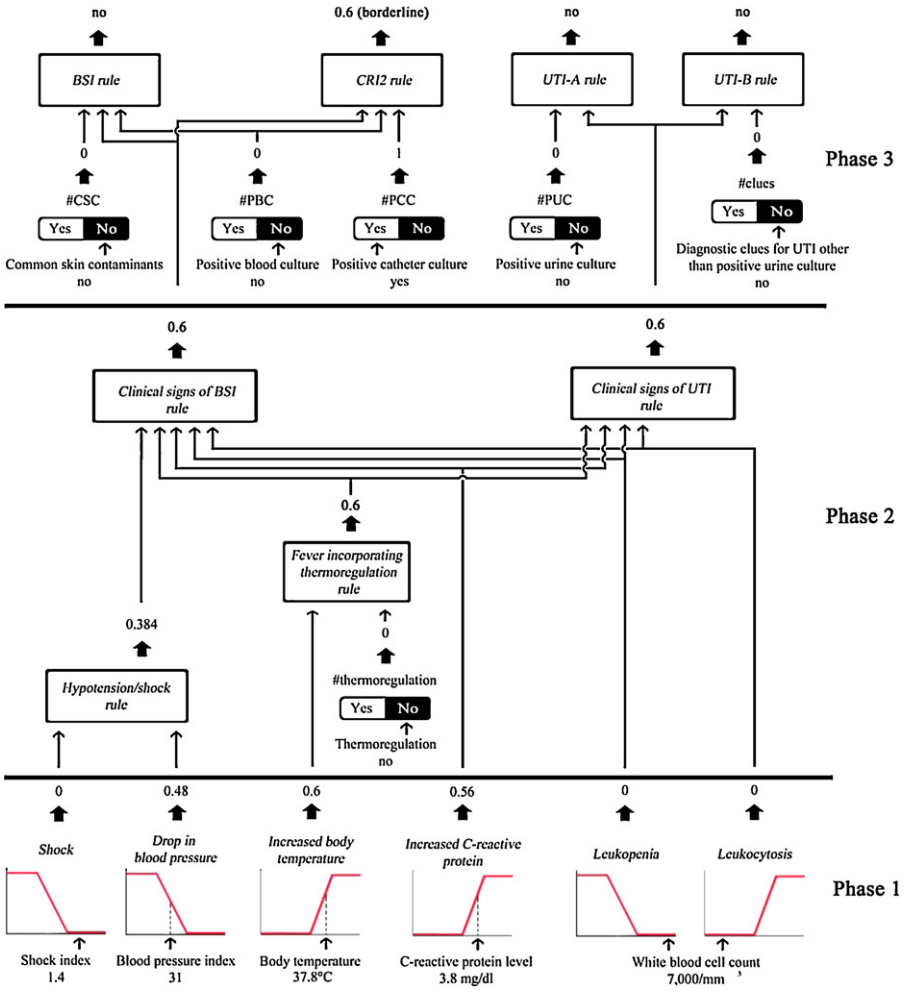

Fig. 2. The knowledge base processing workflow using the data of *Patient X*. In phase 1, values for the fuzzy sets are calculated using the available measured raw patient data. In phase 2, the resulting output values are used together with crisp clinical data to calculate values for more abstract clinical concepts. Finally, in phase 3 the output values for these clinical concepts are used together with crisp microbiology data to calculate the values for the top-level HAI rules.

Table 5
Contingency table for the general group of infection-related concepts relevant to the detection of all types of HAIs.

Clinical concept		All top-level HAIs		*p*-value[a]
		Borderline	Normal	
Increased body temperature	Borderline	115	2,687	0.33
	Normal	1,245	30,516	
Increased C-reactive protein	Borderline	484	4,894	<0.001
	Normal	876	12,102	
Leukopenia	Borderline	55	529	<0.001
	Normal	1,305	39,723	
Leukocytosis	Borderline	71	1,389	0.41
	Normal	1,289	26,094	
Fever incorporating thermoregulation	Borderline	1,150	26,762	<0.001
	Normal	210	6,441	

Note: HAI, healthcare-associated infection.

[a] A *p*-value for a clinical concept indicates the ability of the fuzzy set or rule representing that concept to separate borderline infection cases from normal ones. A $p < 0.05$ means that when a top-level HAI rule associated with the clinical concept was classified as borderline, the concept was also classified significantly more often as borderline rather than normal.

For the assessment of fuzzy sets and fuzzy rules, we divided infection-related concepts into three groups. The first was the general group of infection-related concepts, which were associated with all monitored HAIs (Table 5). The second was the BSI/CRI2-specific group, comprising concepts associated with BSI and CRI2 (Table 6). The third was the UTI-specific group, consisting of concepts solely associated with UTI-A and -B (Table 7).

Analysis showed that borderline values for three infection-related clinical concepts did not co-occur significantly more often with borderline values than with normal values for their associated top-level HAI rules. The three concepts were *increased body temperature* ($p = 0.33$), *leukocytosis* ($p = 0.41$) and *shock* ($p = 0.18$). For all other infection-related clinical concepts the co-occurrence was significant.

Table 6
Contingency table for infection-related concepts relevant to the detection of bloodstream infection and catheter-related infection.

Clinical concept		Top-level BSI and CRI2HAIs		*p*-value[a]
		Borderline	Normal	
Shock	Borderline	219	6,852	0.18
	Normal	523	17,706	
Drop in blood pressure	Borderline	108	3,210	<0.001
	Normal	398	18,104	
Hypotension/shock	Borderline	957	33,615	<0.001
	Normal	191	8,619	
Clinical signs of BSI	Borderline	1,139	10,695	<0.001
	Normal	0	199	

Note: HAI, healthcare-associated infection; BSI, bloodstream infection; CRI, central venous catheter-related infection.
[a] see Table 5.

Table 7
Contingency table for infection-related concepts relevant to the detection of the urinary tract infection.

Clinical concept		Top-level UTI-A and -B HAIs		*p*-value[a]
		Borderline	Normal	
Clinical signs of UTI	Borderline	243	9,760	<0.001
	Normal	0	2,243	

Note: HAI, healthcare-associated infection; UTI, urinary tract infection.
[a] see Table 5.

5. Discussion

We used fuzzy set theory and fuzzy logic to identify patient cases with borderline signs of HAIs, and separated these from patients with no signs of HAIs. The identification of borderline infection is clinically relevant because it allows early recognition of incipient infection. Especially in the ICU setting, where patients' vital signs and blood markers are repeatedly assessed and closely monitored, the onset of infection should be identified rapidly [27]. Early recognition is important because it permits early therapeutic or organizational interventions by clinicians as well as infection control personnel, and improves the patient's outcome [28].

The data showed that, for clinical concepts represented by fuzzy sets and especially for higher-level clinical concepts represented by fuzzy rules, the size of the borderline class was significant. Reclassification ratios were sometimes very high (close to 100%). This accurately reflects the difficulties faced by physicians in the process of diagnosis. The determination of a pathological condition is relatively clear for lower-level clinical concepts related to quantitative patient data, but a clear-cut decision is very difficult when all of these concepts have to be combined. This gives rise to inter-rater variability, which is one of the major drawbacks of manual infection surveillance [29].

Despite the frequent occurrence of borderline values for fuzzy sets and rules, very few borderline values were found for HAI definitions. This is due to several factors. First, all of the discussed HAI definitions involve microbiology laboratory results, which are dichotomous by nature (pathogen grown yes/no). As microbiology tests are not ordered in all cases, the data may not be available for all patients. Even when a relevant microorganism is grown, the result is reported only when microbes are found in relevant numbers (according to the ECDC rules, thresholds from 10^2 to 10^5 colony-forming units per ml of sample apply, depending on the type of specimen [22]). Fuzzy sets could also be constructed for microbiology reports with quantitative features, and borderline values could then be determined. The rule definitions for *clinical signs of BSI* and *clinical signs of UTI* is worthy of mention. The ECDC surveillance definitions contain several infection-related criteria not recorded electronically in the VGH hospital information system. To compensate for this lack of data, the aforementioned fuzzy rules are defined as a disjunction over the large majority of, or all infection parameters (cf., Table 2). This is less restrictive than the ECDC definitions which state that, in some cases, the patient should have at least two or more clinical symptoms. As a result, about 75% of the patient days reveal pathological values for these aggregation rules. While this results in the effective detection of infections [12], it also makes the aggregation parameters less specific in the detection of borderline infection.

The study data frequently showed borderline results for *hypotension/shock*, but never showed fully pathological results. *Hypotension/shock* indicates the temporal occurrence of abnormally low blood pressure, which may be a sign of infection, but may be due to other causes as well, such as medication [30]. To model this lack of evidential power, the compatibility for the *hypotension/shock* rule (cf., Table 2) was adjusted mathematically by the numerical modifier 0.8. An alternative would have been to rate observed *hypotension/shock* signs as a pathological indication for the concept only in cases when ICU personnel manually recorded the presence of infection-related hypotension in a patient. This never occurred during the study period.

The data showed that, for the fuzzy sets for *increased body temperature, leukocytosis*, and *shock,* borderline results did not co-occur significantly more often with borderline values than with normal values for their associated top-level HAI rules. This indicates that the transition ranges of these fuzzy sets are too unspecific. The study data show that the average white blood cell count and body temperature for borderline infection cases is higher than it is for normal patients, and the average shock index is lower (data not shown). This suggests that if the transition range of the associated fuzzy sets were made smaller, they would be more specific.

The strengths of the present study are worthy of mention. It is based on more than 5000 patient stays, comprising almost 50,000 patient days. As such, it provided enough data to accurately analyze the frequency distributions of normal, borderline, and pathological values for infection-related clinical concepts. Another strength is the extensive availability of data (more than 99%). Only 446 out of 49,394 patient days recorded by the ICU's patient data management system were missing due to export or registration problems. Another strength is the transparent description of the system. This report provides enough details to recreate a part of the system for others to perform their own experiments with fuzzy sets and logic for the detection of infection. Finally, to the authors' knowledge

there has been no previous discussion on the identification of borderline infection cases, nor on the optimization of fuzzy sets and rules using statistical methods.

The limitations of the study are also noteworthy. First, no gold standard was available to confirm borderline infection cases. As the system is used in clinical routine, data were evaluated prospectively by infection control experts, and a sample of the results was also evaluated retrospectively for correctness. However, the notion of a borderline infection case is to a certain degree arbitrary; it depends on the opinion of the involved expert(s). As such, the establishment of a gold standard for borderline infection cases relies to a great extent on a subjective interpretation of the ECDC rules, which results in even greater inter-rater variability. Another limitation of the study is the small number of borderline results for HAI definitions, which made it unfeasible to study fuzzy sets and rules for individual ECDC-defined infections. Furthermore, since we had to omit relevant subsets of ECDC-defined HAIs, we could only address those infection entities included in the study. The respective patient may, in fact, have had another ECDC-defined infection or even one not covered by ECDC definitions.

Yet, some of the above limitations need further reflection. What is the definition of a gold standard in this context? As described above, a borderline condition is, by nature, off standard. Intelligent IT tools which use fuzzy sets and fuzzy logic enable experts to judge and even augment their own expertise. Possibly a "gold standard" for borderline conditions could only be established by the use of intelligent IT tools such as those employed in the present study—merely because of the quantity and precision of the required data and the lack of manpower to retrieve such a body of data.

We focused on compatibility between patient data and ECDC-defined rules for HAIs produced by an electronic HAI monitoring system, thereby supporting degrees of compatibility and the identification of an additional borderline patient class rather than a traditional dichotomous classification mechanism. Other studies have determined the effects of various infection risk factors in electronic detection of HAIs [1,2]. A substantial number of electronic HAI detection systems have been created to detect one or more types of HAIs [5,31–33]. However, all systems focus on the detection of definite HAI cases and disregard borderline infection cases. This limits the usefulness of the systems for the detection and prevention of infection because early identification of the signs and symptoms of HAIs may be useful to predict the recurrence of CRI episodes [9].

Fuzzy logic is not the sole means of identifying borderline infection cases. Several authors have used regression analysis methods to detect a variety of HAIs [1,2,34]. These methods employ a probabilistic model to predict the occurrence of an infection, and then define a probability threshold for which the system most accurately predicts infection. It is possible to define a second, lower probability threshold, thereby creating a transition range. Analogous to our method, all patient cases identified within this range could then be considered borderline infection cases.

6. Conclusion

In the present study, we separated borderline infection cases from patients without signs of infection. With the fuzzy sets and fuzzy rules contained in its knowledge base, the Moni-ICU program showed a sizable borderline class for infection-related concepts. However, just a few cases of borderline infection were noted for the top-level HAI definitions. Assessment of the fuzzy sets and rules showed that, especially for *clinical signs of UTI* and *clinical signs of BSI*, borderline indications appeared significantly more often for borderline HAI cases. Based on the results of the study, clinical knowledge engineers and infection control experts are able to tune the knowledge base more accurately to optimize case recognition for both definite and borderline infection cases.

In the future, we aim to utilize fuzzy sets and rules, and degrees of compatibility in general, for a variety of purposes. First, the study showed that p values were very small (<0.001) for several infection-related concepts, which indicates that fuzzy sets for these concepts could be adapted to include a wider range of input values in the transition range. Based on these modifications, we might see more borderline classifications for infection-related concepts and HAI definitions. As the numbers grow, we might be able to divide the borderline class into subclasses, which may be useful for recognizing patterns and generating clinical alerts.

References

[1] van Mourik MS, Groenwold RH, Berkelbach van der Sprenkel JW, van Solinge WW, Troelstra A, Bonten MJ. Automated detection of external ventricular and lumbar drain-related meningitis using laboratory and microbiology results and medication data. PLoS One 2011;6(8):e22846.
[2] Chang YJ, Yeh ML, Li YC, Hsu CY, Lin CC, Hsu MS, et al. Predicting hospital-acquired infections by scoring system with simple parameters. PLoS One 2011;6(8):e23137.
[3] Bolon MK, Hooper D, Stevenson KB, Greenbaum M, Olsen MA, Herwaldt L, et al. Improved surveillance for surgical site infections after orthopedic implantation procedures: extending applications for automated data. Clin Infect Dis 2009;48(9):1223–9.
[4] Sherman ER, Heydon KH, John KHS, Teszner E, Rettig SL, Alexander SK, et al. Administrative data fail to accurately identify cases of healthcare-associated infection. Infect Control Hosp Epidemiol 2006;27(4):332–7.
[5] Trick WE, Zagorski BM, Tokars JI, Vernon MO, Welbel SF, Wisniewski MF, et al. Computer algorithms to detect bloodstream infections. Emerg Infect Dis 2004;10(9):1612–20.
[6] Klompas M, Kleinman K, Platt R. Development of an algorithm for surveillance of ventilator-associated pneumonia with electronic data and comparison of algorithm results with clinician diagnoses. Infect. Control Hosp Epidemiol 2008;29(1):31–7.
[7] JT Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh.George JK, Bo Y, editors. River Edge, NJ, USA: World Scientific Publishing Co., Inc.; 1996.
[8] de Bruin JS, Seeling W, Schuh C. Data use and effectiveness in electronic surveillance of healthcare associated infections in the 21st century: a systematic review. J Am Med Inform Assoc 2014;21(5):942–51.
[9] de Bruin JS, Blacky A, Adlassnig K-P. Assessing the clinical uses of fuzzy detection results in the automated detection of CVC-related infections: a preliminary report. In: Mantas J, Andersen SK, Mazzoleni MC, Blobel B, Quaglini S, Moen A, editors. Stud Health Technol Inform, vol. 180. Amsterdam: IOS Press; 2012. p. 579–83.
[10] Adlassnig K-P, Blacky A, Koller W. Fuzzy-based nosocomial infection control. In: Nikravesh M, Kacprzyk J, Zadeh L, editors. Forging New Frontiers: Fuzzy Pioneers II, vol. 218. Berlin: Springer; 2008. p. 343–9.
[11] Adlassnig K-P, Berger A, Koller W, Blacky A, Mandl H, Unterasinger L, Rappelsberger A. Healthcare-associated infection surveillance and bedside alerts. In: Hörbst A, Hayn D, Schreier G, Ammenwerth E, editors. eHealth2014—Health Informatics Meets eHealth, Outcomes Research: The Benefit of Health-IT, Studies in Health Technology and Informatics, vol. 198. Amsterdam: IOS Press; 2014. p. 71–8.
[12] de Bruin JS, Adlassnig K-P, Blacky A, Mandl H, Fehre K, Koller W. Effectiveness of an automated surveillance system for intensive care unit-acquired infections. J Am Med Inform Assoc 2013;20(2):369–72.
[13] Zadeh LA. Fuzzy sets. Inf Control 1965;8(3):338–53.
[14] Adlassnig KP. Uniform representation of vagueness and imprecision in patientís medical findings using fuzzy sets. In: Trappl R, editor. Cybernetics and Systemsí88. Dordrecht: Kluwer Academic Publishers; 1988. p. 685–92.
[15] Adlassnig K-P, Blacky A, Mandl H, Rappelsberger A, Koller W. Fuzziness in healthcare-associated infection monitoring and surveillance. In: Winston A, Melek W, Hall P, McKenzie S, Gibbs M, Adamson G, editors. Proceedings of the 2014 IEEE Conference on Norbert Wiener in the 21st Century–Driving Technology's Future. Piscataway: Institute of Electrical and Electronics Engineers (IEEE); 2014, 86.pdf.
[16] Zadeh LA. Fuzzy aqlgorithms. Inf Control 1968;12(2):94–102.
[17] Klement EP, Mesiar R, Pap E. Triangular Norms. Dordrecht, Boston: Kluwer Academic Publishers; 2000.
[18] Adlassnig K-P. A survey on medical diagnosis and fuzzy subsets. In: Gupta MM, Sanchez E, editors. Approximate Reasoning in Decision Analysis. Amsterdam: North Holland Publishing Company; 1982. p. 203–17.
[19] Steimann F. On the use and usefulness of fuzzy sets in medical AI. Artif Intell Med 2001;21(1–3):131–7.
[20] Abbod MF, Keyserlingk DGv, Linkens DA, Mahfouf M. Survey of utilisation of fuzzy technology in medicine and healthcare. Fuzzy Sets Syst 2001;120(2):331–49.

[21] Mahfouf M, Abbod MF, Linkens DA. A survey of fuzzy logic monitoring and control utilisation in medicine. Artif Intell Med 2001;21(1–3):27–42.
[22] European Centre for Disease Prevention and Control (ECDC). HAIICU Protocol v1.01 Standard and Light; 2010. http://www.ecdc.europa.eu/en/aboutus/calls/Procurement%20Related%20Documents/5_ECDC_HAIICU_protocol_v1_1.pdf (accessed 11.04.16).
[23] Health Level Seven Inc. The Arden Syntax for Medical Logic Systems Version 2.7; 2008. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=2 (accessed 11.04.16).
[24] Hripcsak G. Writing arden syntax medical logic modules. Comput Biol Med 1994;24(5):331–63.
[25] Centers for Disease Control and Prevention (CDC). 2016 NHSN Patient Safety Component Manual; 2016. http://www.cdc.gov/nhsn/pdfs/pscmanual/pcsmanual_current.pdf (accessed 11.04.16).
[26] Nationales Referenzzentrum für Surveillance von nosokomialen Infektionen. Surveillance nosokomialer Infektionen auf Intensivstationen; 2015. http://www.nrz-hygiene.de/fileadmin/nrz/download/ITS-KISS-InfSurv_Protokoll_v20151028.pdf (accessed 11.04.16).
[27] Charles PE, Kus E, Aho S, Prin S, Doise JM, Olsson NO, et al. Serum procalcitonin for the early recognition of nosocomial infection in the critically ill patients: a preliminary report. BMC Infect Dis 2009;9:49.
[28] Harbarth S, Garbino J, Pugin J, Romand JA, Lew D, Pittet D. Inappropriate initial antimicrobial therapy and its effect on survival in a clinical trial of immunomodulating therapy for severe sepsis. Am J Med 2003;115(7):529–35.
[29] Haley RW, Culver DH, White JW, Morgan WM, Emori TG, Munn VP, et al. The efficacy of infection surveillance and control programs in preventing nosocomial infections in US hospitals. Am J Epidemiol 1985;121(2):182–205.
[30] National Institute of Health, What is Hypotension? http://www.nhlbi.nih.gov/health/health-topics/topics/hyp/ (accessed 11.04.16).
[31] Bouzbid S, Gicquel Q, Gerbier S, Chomarat M, Pradat E, Fabry J, et al. Automated detection of nosocomial infections: evaluation of different strategies in an intensive care unit 2000–2006. J Hosp Infect 2011;79(1):38–43.
[32] Leth RA, Møller JK. Surveillance of hospital-acquired infections based on electronic hospital registries. J Hosp Infect 2006;62(1):71–9.
[33] Woeltje KF, Butler AM, Goris AJ, Tutlam NT, Doherty JA, Westover MB, et al. Automated surveillance for central line-associated bloodstream infection in intensive care units. Infect Control Hosp Epidemiol 2008;29(9):842–6.
[34] Yokoe DS, Christiansen CL, Johnson R, Sands KE, Livingston J, Shtatland ES, et al. Epidemiology of and surveillance for postpartum infections. Emerg Infect Dis 2001;7(5):837–41.