


Article

Discriminating Bacterial Infection from Other Causes of Fever Using Body Temperature Entropy Analysis

Borja Vargas ^{1,*}, David Cuesta-Frau ², Paula González-López ¹, María-José Fernández-Cotarelo ^{1,3}, Óscar Vázquez-Gómez ^{1,3}, Ana Colás ⁴ and Manuel Varela ¹

¹ Department of Internal Medicine, Hospital Universitario de Móstoles, 28935 Mostoles, Spain; paula.gonzalez.lopez@salud.madrid.org (P.G.-L.); mfcotarelo@salud.madrid.org (M.-J.F.-C.); ovazquez@salud.madrid.org (Ó.V.-G.); manuel.varela@salud.madrid.org (M.V.)

² Technological Institute of Informatics, Universitat Politècnica de València, Alcoi Campus, 03801 Alcoi, Spain; dcuesta@disca.upv.es

³ Faculty of Health Sciences, Universidad Rey Juan Carlos, 28922 Alcorcon, Spain

⁴ Department of Internal Medicine, Hospital Universitario 12 de Octubre, 28041 Madrid, Spain; anamaria.colas@salud.madrid.org

* Correspondence: borja.vargas@salud.madrid.org; Tel.: +34-669-799-079

Abstract: Body temperature is usually employed in clinical practice by strict binary thresholding, aiming to classify patients as having fever or not. In the last years, other approaches based on the continuous analysis of body temperature time series have emerged. These are not only based on absolute thresholds but also on patterns and temporal dynamics of these time series, thus providing promising tools for early diagnosis. The present study applies three time series entropy calculation methods (Slope Entropy, Approximate Entropy, and Sample Entropy) to body temperature records of patients with bacterial infections and other causes of fever in search of possible differences that could be exploited for automatic classification. In the comparative analysis, Slope Entropy proved to be a stable and robust method that could bring higher sensitivity to the realm of entropy tools applied in this context of clinical thermometry. This method was able to find statistically significant differences between the two classes analyzed in all experiments, with sensitivity and specificity above 70% in most cases.

Keywords: time series; body temperature; Slope Entropy; Approximate Entropy; Sample Entropy; classification; fever



Citation: Vargas, B.; Cuesta-Frau, D.; González-López, P.; Fernández-Cotarelo, M.-J.; Vázquez-Gómez, Ó.; Colás, A.; Varela, M. Discriminating Bacterial Infection from Other Causes of Fever Using Body Temperature Entropy Analysis. *Entropy* **2022**, *24*, 510. <https://doi.org/10.3390/e24040510>

Academic Editor: Roberto Hornero

Received: 23 February 2022

Accepted: 2 April 2022

Published: 5 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Body temperature is a key clinical parameter. It is usually assessed once per shift in hospital wards and has always been considered a hallmark of infectious diseases. However, the values obtained with the standard measurements are interpreted dichotomously: either the patient has a fever or is afebrile.

Body temperature assessment is also highly dependent on the method of measurement. Central methods are accurate and reliable (pulmonary artery catheter, urinary bladder, esophagus) but they are not suitable in most clinical scenarios. Tympanic temperature is often used as a replacement for central temperature because values are close, and it is more convenient and less invasive [1]. Peripheral temperature can be assessed in different anatomical locations (mouth, armpit). Despite not being as accurate [2], peripheral measurements are the standard procedure in clinical practice.

Furthermore, the definition of fever is arguably flawed, as it depends on many factors such as age, gender, circadian rhythms, or underlying conditions [3–5]. As a matter of fact, there is no universal threshold for fever, as a wide range of temperatures has been shown in individuals considered healthy [6,7]. Some efforts to standardize the *normal* body temperature range have been carried out in the past [7] but they have not been transferred into clinical practice.

Traditionally, attempts have been made to find clinical differences in the patterns of fever caused by infectious diseases (malaria, tuberculosis, typhoid fever) [8]. Nevertheless, none of these approaches are sufficient to make clinical decisions [8,9]. Furthermore, a wide spectrum of noninfectious conditions can also induce the synthesis and release of pyrogenic cytokines and eventually cause fever [10].

Since body temperature regulation is a dynamical process, by obtaining just two or three measurements per day, a wealth of information is lost. However, some devices are available to obtain high-frequency measurements of body temperature. Body temperature monitoring has been proven useful in certain clinical scenarios when more frequent measurements (associated or not to alarm settings) entail earlier recognition of fever [11–14]. Moreover, temperature monitoring devices enable the registry of temperature time series. This allows its use as a continuous variable, instead of a series of isolated values [15].

Similar to many other biological systems, thermoregulation can be considered a complex process, and might therefore be analyzed under the scope of nonlinear dynamics. Complexity metrics have previously been applied to other biological variables [16]. It has been widely demonstrated that changes in complexity of biological signals are associated with damage to or degradation of the system [17–23].

In this context, entropy statistics could be of clear interest to unveil certain characteristics of the thermoregulation process and, perhaps, the underlying cause of fever. In previous works, we have already demonstrated the feasibility of this approach. For example, in [24], we described a method based on the entropy statistic Slope Entropy (SlpEn) [25] to distinguish between body temperature time series from malaria and dengue patients. The achieved accuracy was up to 90% correctly classified records with a single numerical feature computed for each one. In another study [26], a different entropy method, Sample Entropy (SampEn) [27], was used with the same purpose of distinguishing among body temperature time series coming from infectious diseases, tuberculosis, nontuberculosis, and dengue fever patients. The global accuracy achieved was close to 70%.

Other works have used a combination of features; this is the case for the work described in [28], which used temperature temporal patterns to detect tuberculosis. In [29], the authors used a more sophisticated approach using the Fourier transform, entropy, energy, power, and a set of additional coefficients to train a quadratic support vector machine to carry out the classification of tuberculosis, intracellular bacterial infections, dengue, and inflammatory and neoplastic diseases temperature time series.

The goal of this study is to assess if patients with bacterial infections have significant changes in the entropy of their body temperature compared with patients with other infections or other causes of fever. As the entropy statistic for the analysis, we chose SlpEn for its good performance in previous studies [24,30,31]. For comparative purposes, we included more widely used methods such as Approximate Entropy (ApEn) [32] and SampEn, which have been successfully used in a myriad of similar biosignal classification works [33–39].

2. Materials and Methods

2.1. SlpEn

The recently proposed time series entropy measure termed Slope Entropy (SlpEn) [25] can achieve high classification accuracy using a diverse set of records [24,25,30]. Despite its short life, it has already been implemented in scientific software tools such as EntropyHub (<https://github.com/MattWillFlood/EntropyHub.jl>, accessed on 15 February 2022) and CEPS (Complexity and Entropy in Physiological Signals) [40].

The first step of SlpEn computation is extraction from an input time series $\mathbf{x} = \{x_0, x_1, \dots, x_{N-1}\}$ of a set of consecutive overlapping subsequences of length m , commencing at sample i , $\mathbf{x}_i = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$, $0 \leq i < N - m + 1$ (m being the embedded dimension variable and n the total length of the time series, with $m \ll N$). Each of the n_m extracted subsequences, \mathbf{x}_i , can then be transformed into a new one of length $m - 1$ by computing and storing the differences between each pair of consecutive samples in the subsequence, namely, $\mathbf{y}_i = \{x_i - x_{i+1}, x_{i+1} - x_{i+2}, \dots, x_{i+m-2} - x_{i+m-1}\}$.

Using, in its basic configuration [25], 5 different symbols from an alphabet—for example +2, +1, 0, −1, −2—the differences obtained are represented by these symbols instead, according to two input thresholds, δ and γ , and the expressions described in [25]. Further details of SlpEn implementation and examples can be found in [24,25]. A software library using this method is also described in [40].

In addition to SlpEn, two other entropy methods—ApEn and SampEn—were applied to the time series in order to assess the hypothesized improvement in classification accuracy that SlpEn could bring to the analysis. Although these methods have been used extensively, and they are characterized and described in great detail in a number of publications [41–46], they are depicted for completeness in the next two subsections.

2.2. Approximate Entropy

ApEn [32] is also based on extracting subsequences of length m from the input time series, $\mathbf{x}_i = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$, as for SlpEn. Then, a distance is computed between every subsequence and a fixed reference \mathbf{x}_j , $d_{ij} = \max(|x_{i+k} - x_{j+k}|)$, with $0 \leq k \leq m - 1$.

If the number of comparisons falling below a predefined threshold r —termed matches, $d_{ij} < r$ —is computed for two consecutive embedded dimensions (m and $m \leftarrow m + 1$), two counters can be defined as $B_i(r)$, number of j so that $d_{ij}^m \leq r$, and $A_i(r)$, number of j so that $d_{ij}^{m+1} \leq r$, with $0 \leq j < N - m + 1$.

Computing the averages of these counters, $B_i^m(r) = \frac{1}{N - m + 1} B_i(r)$ and $A_i^m(r) = \frac{1}{N - m} A_i(r)$, the main ApEn variables are calculated as $\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \log B_i^m(r)$ and $\phi^{m+1}(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} \log A_i^m(r)$, from which the result of ApEn can be finally obtained as $\text{ApEn}(m, r, N) = [\phi^m(r) - \phi^{m+1}(r)]$.

2.3. Sample Entropy

The first steps of the SampEn algorithm [35] are the same as for ApEn. However, when counting the matches, subsequences are not compared with themselves, formally $0 \leq j < N - m + 1$, with $j \neq i$.

Then, the statistics are now $B^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r)$ and $A^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A_i^m(r)$, from which SampEn is computed as $\text{SampEn}(m, r, N) = -\log \left[\frac{A^m(r)}{B^m(r)} \right]$.

2.4. Experimental Dataset

The study was conducted at Hospital Universitario de Móstoles (Madrid, Spain). Patients older than 18 years old admitted to the general Internal Medicine ward presenting with fever at admission and/or suspected infectious disease were considered suitable for inclusion. Pregnancy and inability to cooperate with the monitoring process were considered exclusion criteria.

Temperature values were obtained through a probe (Truer Medical, Inc., Orange, CA, USA) placed in the external auditory canal (EAC), after otoscopy to check the integrity of the tympanic membrane. Data from the EAC were used as surrogates of central temperature [1]. The probe was wired to a Holter device (TherCom, Innovatec) that registered one measurement per minute. When feasible, the monitoring process was performed in real-time. Otherwise, data were stored in the device and downloaded later for analysis. The aim was to perform 24-h recordings, but in some cases, the process was stopped earlier due to poor compliance of the patient, displacement of the probes for long periods (preventing the proper recording of data), or abnormally low temperatures, suggesting that measurements were clearly inaccurate.

Patients were classified into two categories concerning diagnosis: bacterial infection (confirmed or suspected) or others. The latter included patients with nonbacterial infections (viral, fungal, etc.) or with fever caused by inflammatory diseases, cancer, or fever of unknown cause (when bacterial infection was deemed to be excluded).

Temperature time series were processed by visual inspection. In some cases, the beginning and/or the end of the recording were trimmed to ensure the stability of the signal. Disconnections of at most 5 measurements were repaired through linear interpolation. For longer disconnections, the segment was removed, provided the remaining interval was clean.

The experimental dataset contained 10 body temperature time series of patients with confirmed bacterial infection and 13 from patients with other causes of fever. The lengths of the time series are shown in Table 1. In all cases, different fixed time series lengths were used to assess the classification accuracy of each metric and ensure length equality: 500, 600, 700, 800, 900, 1000, and 1100. Those time series below the cut-off length were discarded in that specific experiment.

Table 1. Original lengths of the body time series used in the experiments. Length is defined in terms of number of samples, taking into account that the sampling frequency was one sample per minute.

	1	2	3	4	5	6	7	8	9	10	11	12	13
Bacterial infection	936	1231	1154	1279	1443	1134	680	710	586	1117	–	–	–
Other causes of fever	1284	1468	1427	1444	1295	913	1105	1017	830	537	1121	859	934

Figure 1 depicts two body temperature records from this database, with one from each class (fever caused by a bacterial infection, and fever caused by a different clinical condition).

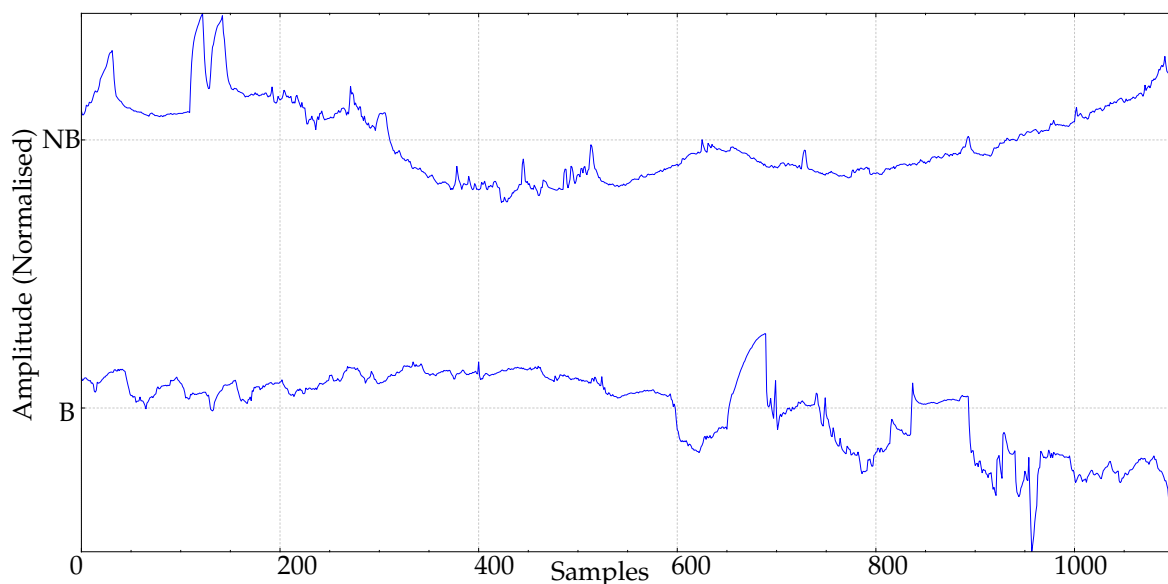


Figure 1. Example of body temperature records from the experimental database. B: Bacterial infection. NB: Nonbacterial cause of fever.

3. Experiments and Results

All the experimental time series were processed using the three entropy calculation methods described previously: SlpEn, ApEn, and SampEn. They were also cut short to the lengths stated above. The remaining central part was used for the analysis, as it would theoretically be the most stable segment (thermal equilibrium reached, probe still in place). The entropy result was used as the classification feature applying Sensitivity (Se) and

Specificity (Sp) [47], with a threshold obtained from the corresponding ROC curve (closest point to (0,1); an example is shown in Figure 2) [48–50]. The statistical significance was assessed using the Wilcoxon–Mann–Whitney test [51], with $\alpha = 0.05$. Input parameters were varied in the range $m \in [3, 9]$ and $\gamma, r \in [0.10, 0.90]$. For SlpEn, δ was kept constant at $\delta = 0.001$. Time series were normalized for zero mean and unit standard deviation. The stationarity of the input time series was assessed by computing the standard deviation for each consecutive 50-sample window, yielding fairly similar values.

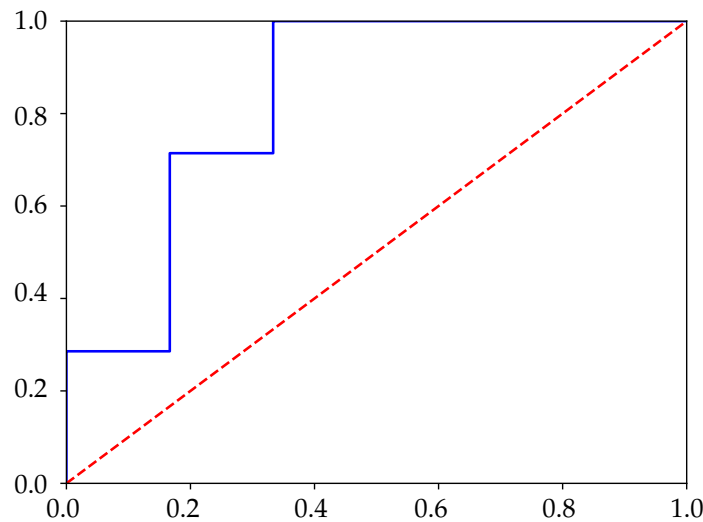


Figure 2. Example of ROC curve using Slope Entropy (SlpEn), $m = 3$, and $\gamma = 0.20$.

Table 2 shows the results for lengths $N = 500, 600$, and 700 . For $N = 500$, SlpEn achieved good classification accuracy for $m = 3$ and in the γ region of 0.15 – 0.30 . Neither ApEn nor SampEn reached significance for any combination of their input parameters m and r , after a grid search for m between 3 and 9 , and r between 0.10 and 0.90 in 0.05 steps. It is important to note that these methods are very sensitive to length, and $N = 500$ is arguably too short for them.

For $N = 600$ and $N = 700$ the results were similar. ApEn and SampEn did not discriminate, SlpEn remained stable in the same region of $m = 3$ and $\gamma = 0.15$ – 0.25 , but other parameter configurations in the same region of γ with values of m such as $4, 6$, and 7 also reached discriminatory power.

Table 2. Experiment results for lengths $N = 500, 600, 700$ using SlpEn, Approximate Entropy (ApEn), and Sample Entropy (SampEn). Parameter grid search for m , between 3 and 9 , and r and γ , between 0.10 and 0.90 in 0.05 steps. The values of the input parameters are included as (m, r) or (m, γ) for cases when $p < 0.05$ after the grid search. Otherwise, no combination provided significant results, represented by $--$. Statistical significance was only reached by SlpEn.

	N = 500			N = 600			N = 700					
	Parameters	p	Se	Sp	Parameters	p	Se	Sp	Parameters	p	Se	Sp
SlpEn	(m = 3, $\gamma = 0.15$)	0.0255	0.90	0.76	(m = 3, $\gamma = 0.15$)	0.0190	0.88	0.75	(m = 3, $\gamma = 0.15$)	0.0448	0.87	0.75
	(3, 0.20)	0.0407	0.69	0.61	(3, 0.20)	0.0330	0.77	0.75	(3, 0.20)	0.0448	1	0.66
	(3, 0.25)	0.0299	0.80	0.76	(3, 0.25)	0.0466	1	1	(4, 0.15)	0.0448	0.87	0.66
	(3, 0.30)	0.0545	1	0.61	(4, 0.15)	0.0330	1	0.66	(6, 0.10)	0.0307	0.66	1
	--	p > 0.05	--	--	(4, 0.20)	0.0466	0.77	0.66	(7, 0.20)	0.0372	0.75	0.87
ApEn	--	p > 0.05	--	--	--	p > 0.05	--	--	--	p > 0.05	--	--
SampEn	--	p > 0.05	--	--	--	p > 0.05	--	--	--	p > 0.05	--	--

Table 3 displays the results for lengths $N = 800$ and 900 . For $N = 800$, the trend is the same as in Table 2. SlpEn is able to find differences in the vicinity of $m = 3$ and $\gamma = 0.20$,

but ApEn and SampEn are unable to yield any statistically significant classification. For $N = 900$, the number of parameter combinations for SlpEn increases, with γ fairly stable in the same region around 0.20, for almost any m value except 5. Additionally, ApEn is also significant in the region around $m = 3$ and $r = 0.20$.

Table 3. Experiment results for lengths $N = 800, 900$ using SlpEn, ApEn, and SampEn. Parameter grid search for m , between 3 and 9, and r and γ , between 0.10 and 0.90 in 0.05 steps. The values of the input parameters are included as (m, r) or (m, γ) for cases when $p < 0.05$ after the grid search. Otherwise, no combination provided significant results, represented by $--$. Statistical significance was reached by SlpEn and ApEn.

$N = 800$				$N = 900$			
Parameters	p	Se	Sp	Parameters	p	Se	Sp
$(m = 3, \gamma = 0.15)$	0.0179	1	0.75	$(m = 3, \gamma = 0.10)$	0.0247	0.71	0.80
$(3, 0.20)$	0.0425	0.85	0.66	$(3, 0.15)$	0.0191	0.71	0.90
$(4, 0.15)$	0.0346	1	0.66	$(3, 0.20)$	0.0317	0.71	0.90
--	$p > 0.05$	--	--	$(3, 0.25)$	0.0404	0.71	0.90
--	$p > 0.05$	--	--	$(4, 0.10)$	0.0146	0.71	0.90
--	$p > 0.05$	--	--	$(4, 0.15)$	0.0247	0.71	0.90
SlpEn	--	$p > 0.05$	--	$(4, 0.20)$	0.0317	0.71	0.90
--	$p > 0.05$	--	--	$(4, 0.35)$	0.0404	0.71	0.70
--	$p > 0.05$	--	--	$(6, 0.10)$	0.0191	0.80	0.85
--	$p > 0.05$	--	--	$(7, 0.10)$	0.0247	0.70	0.85
--	$p > 0.05$	--	--	$(7, 0.15)$	0.0317	0.70	0.85
--	$p > 0.05$	--	--	$(7, 0.20)$	0.0404	0.70	0.85
--	$p > 0.05$	--	--	$(8, 0.10)$	0.0317	0.80	0.85
--	$p > 0.05$	--	--	$(8, 0.15)$	0.0191	0.90	0.71
--	$p > 0.05$	--	--	$(8, 0.20)$	0.0317	0.90	0.71
--	$p > 0.05$	--	--	$(9, 0.15)$	0.0317	0.70	0.85
ApEn	--	$p > 0.05$	--	$(m = 3, r = 0.20)$	0.0317	0.85	0.7
--	$p > 0.05$	--	--	$(3, 0.25)$	0.0317	0.85	0.7
--	$p > 0.05$	--	--	$(4, 0.20)$	0.0404	0.85	0.7
--	$p > 0.05$	--	--	$(4, 0.25)$	0.0247	1	0.7
SampEn	--	$p > 0.05$	--	--	$p > 0.05$	--	--

Finally, the results for $N = 1000$ and $N = 1100$ are shown in Table 4. No more lengths were tested since not enough time series would be available if $N > 1100$. For both cases, the number of significant combinations increased significantly, with SlpEn certainly stable in the same regions as for other N values, and even ApEn and SampEn reaching significance for $N = 1100$. Figure 3 shows a plot of results for $N = 1100, \gamma = r = 0.20, m = 3$ and for the three methods tested.

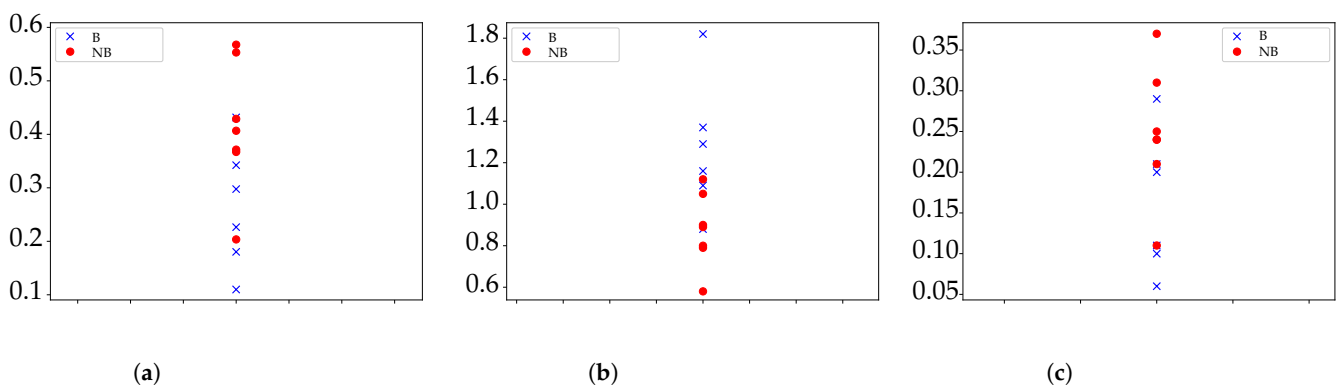


Figure 3. Example of graphical results for each method tested with $N = 1100$. SlpEn results have been inverted and rescaled for better visualization. (a) Results for ApEn with $r = 0.2$ and $m = 3$. (b) Results for SlpEn with $\gamma = 0.2$ and $m = 3$ (In absolute value and normalized by 100). (c) Results for SampEn with $r = 0.2$ and $m = 3$.

Table 4. Experiment results for lengths $N = 1000, 1100$ using SIpEn, ApEn, and SampEn. Parameter grid search for m , between 3 and 9, and r and γ , between 0.10 and 0.90 in 0.05 steps. The values of the input parameters are included as (m, r) or (m, γ) for cases when $p < 0.05$ after the grid search. Otherwise, no combination provided significant results, represented by $--$. Statistical significance was reached by all methods in some cases.

	$N = 1000$				$N = 1100$			
	Parameters	p	Se	Sp	Parameters	p	Se	Sp
SlpEn	$(m = 3, \gamma = 0.1)$	0.0388	0.83	0.87	$(m = 3, \gamma = 0.1)$	0.0222	0.83	0.85
	$(3, 0.15)$	0.0388	0.83	0.87	$(3, 0.15)$	0.0222	0.83	0.85
	$(4, 0.10)$	0.0388	0.83	0.75	$(3, 0.20)$	0.0222	0.83	0.85
	$(4, 0.15)$	0.0281	1	0.75	$(3, 0.25)$	0.0151	0.83	0.85
	$(8, 0.10)$	0.0388	0.75	0.83	$(3, 0.30)$	0.0222	0.83	0.85
	$(8, 0.15)$	0.0281	0.75	1	$(3, 0.35)$	0.0101	1	0.85
	$(9, 0.10)$	0.0388	0.75	0.83	$(3, 0.40)$	0.0321	0.83	0.71
	$(9, 0.15)$	0.0098	0.87	0.83	$(4, 0.10-0.40)$	0.0101	0.83	0.85
	$(9, 0.20)$	0.0388	0.87	0.66	$(6, 0.10)$	0.0222	0.85	0.83
	$(9, 0.25)$	0.0281	0.75	0.83	$(6, 0.15)$	0.0222	0.85	1
	$(9, 0.35)$	0.0388	0.87	0.83	$(7, 0.10)$	0.0151	0.85	1
	--	--	--	--	$(7, 0.15)$	0.0222	0.85	0.83
	--	--	--	--	$(8, 0.10-0.35)$	0.0321	0.85	0.83
	--	--	--	--	$(9, 0.15-0.35)$	0.0151	0.71	1
ApEn	--	--	--	$(m = 9, r = 0.1)$	0.0151	0.85	0.83	
	--	--	--	$(10, 0.1)$	0.0321	0.85	0.83	
SampEn	--	--	--	$(m = 3, r = 0.1-0.6)$	0.0455	0.83	0.85	

4. Discussion

The experiments explored the capability of SIpEn, ApEn and SampEn to distinguish between two classes of body temperature records: time series from patients with bacterial infection and time series from patients also with fever but due to other causes. The experimental set available enabled a study using lengths from 500 up to 1100 samples.

For all these lengths, SIpEn was able to find significant differences when the input parameters were $m = 3$ and $\gamma = 0.15-0.25$, with additional m values available depending on N values. This illustrates the fact that SIpEn is fairly stable and robust, as also demonstrated in other studies based on this recent method [30,31,52].

The results obtained using ApEn were significant only for $N = 900$ (Table 3) and $N = 1100$ (Table 4). This is in accordance with the reported high sensitivity of ApEn to the length of the input time series [44]. One of the most popular guidelines for this minimum length using ApEn is $N \geq 10^m$ [53], which translates in this case to $N \geq 1000$, in agreement with the results in this study. In order to illustrate how the ApEn statistics were computed for the most unfavorable case, $N = 500$, the percentage of estimated probabilities on at least 10 matches, weak criterion, was 80.17 ± 10.37 , and on at least 100 matches, strong criterion, was 6.52 ± 13.01 [54].

SampEn only achieved significance for $N = 1100$. Although SampEn is known to be sensitive to input time series length, it is usually claimed to be more robust in this regard than ApEn. However, in other cases, we have also found that ApEn performed better than SampEn in classification tasks such as in [38].

There is also an association between time series length and classification performance in terms of specificity and sensitivity [47]. For shorter time series, there are some cases where these metrics achieve values below 0.7. As the length increases, the performance improves, with more values in the vicinity of 0.85, arguably very high for a biomedical signal classification application.

Therefore, for lengths shorter than 1000 samples, imposed by the operational difficulties linked to obtaining long-term body temperature records, SIpEn seems a good choice to find differences between the record classes present in the experimental dataset. If longer series were available, other methods such as ApEn and SampEn could also be applied, which remains to be further studied.

From a clinical perspective, the results of this work suggest that patients admitted to the hospital with a diagnosis of bacterial infection had a misregulation of their body temperature, measured with entropy statistics. This is in accordance with findings of a

previous study by our group [55]. The complexity of biosignal time series seems to be an indicator of the integrity and performance of biological systems, and disease usually exhibits low levels of entropy metrics [16,20]. Bacterial infection may entail the development of sepsis, a clinical entity with a very high risk of complications and death, which is unusual with other infections or other causes of fever. In our opinion, the loss of entropy that we have observed in the temperature curves of patients with bacterial infections may be another facet of homeostasis disturbance.

Remarkably, these results were irrespective of the confirmation of fever by the staff nurse (standard measurements), or the maximum temperature obtained. These results suggest that body temperature may supply relevant information, over and above attaining a certain pre-established *febrile threshold*.

In fact, body temperature may provide clues in many clinical aspects, as long as enough information is obtained through continuous monitoring. It has already proved to be useful to assess the prognosis of critically ill patients in the Intensive Care Unit [19,56], to forecast fever peaks [13], and to classify patients according to the cause of fever [26,29].

We are aware of the many limitations of this work. On the one hand, the monitoring system has some issues: the tympanic probe is prone to be displaced, wired probes can be bothersome for the patient, the Holter device needs to be wirelessly connected to a computer in the range of Bluetooth, etc. For these reasons, over half of the patients in the study were excluded from this analysis because the recordings were lost or defective. Several adjustments were carried out during the study to solve or reduce the impact of these issues, such as real-time monitoring through a wireless network and periodic backups of data to keep a copy of the recording in case there was a disconnection. In any case, we acknowledge that the final sample was small and, although significant differences have been found between the two groups, reliability might be limited for this reason.

On the other hand, as it has already been exposed, entropy metrics need clean time series and are in general more informative the longer the data. This has been a limitation in this work and is a common problem for the analysis of biological time series recorded in real life, as many factors may cause artifacts and make data unsuitable for evaluation.

In our opinion, future research should focus on two issues: Firstly, acquisition of long and clean time series. For this purpose, wireless and ergonomic probes that fit properly at the external auditory canal could improve the quality of the recordings and minimize the loss of information. Secondly, obtaining temperature recordings in a wide range of clinical settings—including healthy individuals for comparison—may provide details about physiological processes and may broaden the utility of clinical thermometry to subtler issues than just the identification of fever peaks.

Author Contributions: Project administration and funding acquisition, B.V.; patient recruitment and clinical data analysis, B.V., P.G.-L., M.-J.F.-C. and Ó.V.-G.; time series and statistical analysis, B.V., A.C., D.C.-F. and M.V.; entropy analysis, D.C.-F.; writing—original draft preparation, B.V., A.C., M.V. and D.C.-F. All authors reviewed the initial version of the paper and contributed to the final one. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Instituto de Salud Carlos III, grant number PI17/00856 (Co-funded by European Regional Development Fund, “A way to make Europe”). The APC was funded by Instituto de Salud Carlos III.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Hospital Universitario de Móstoles with protocol code 2017/033 on 26 October 2017.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality reasons.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Uleberg, O.; Eidstuen, S.C.; Vangberg, G.; Skogvoll, E. Temperature measurements in trauma patients: Is the ear the key to the core? *Scand. J. Trauma Resusc. Emerg. Med.* **2015**, *23*, 1–8. [[CrossRef](#)]
2. Niven, D.J.; Gaudet, J.E.; Laupland, K.B.; Mrklas, K.J.; Roberts, D.J.; Stelfox, H.T. Accuracy of peripheral thermometers for estimating temperature: A systematic review and meta-analysis. *Ann. Intern. Med.* **2015**, *163*, 768–777. [[CrossRef](#)] [[PubMed](#)]
3. Speaker, S.L.; Pfoh, E.R.; Pappas, M.A.; Hu, B.; Rothberg, M.B. Oral Temperature of Noninfected Hospitalized Patients. *JAMA* **2021**, *325*, 1899–1901. [[CrossRef](#)] [[PubMed](#)]
4. Mortola, J.P. Gender and the circadian pattern of body temperature in normoxia and hypoxia. *Respir. Physiol. Neurobiol.* **2017**, *245*, 4–12. [[CrossRef](#)] [[PubMed](#)]
5. Roghmann, M.C.; Warner, J.; Mackowiak, P.A. The relationship between age and fever magnitude. *Am. J. Med. Sci.* **2001**, *322*, 68–70. [[CrossRef](#)] [[PubMed](#)]
6. Obermeyer, Z.; Samra, J.K.; Mullainathan, S. Individual differences in normal body temperature: Longitudinal big data analysis of patient records. *BMJ (Online)* **2017**, *359*, 5468. [[CrossRef](#)] [[PubMed](#)]
7. Mackowiak, P.A.; Wasserman, S.S.; Levine, M.M. A Critical Appraisal of 98.6 °F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich. *JAMA J. Am. Med. Assoc.* **1992**, *268*, 1578–1580. [[CrossRef](#)]
8. Ogoina, D. Fever, fever patterns and diseases called ‘fever’—A review. *J. Infect. Public Health* **2011**, *4*, 108–124. [[CrossRef](#)]
9. Musher, D.M.; Fainstein, V.; Young, E.J.; Pruett, T.L. Fever Patterns: Their Lack of Clinical Significance. *Arch. Intern. Med.* **1979**, *139*, 1225–1228. [[CrossRef](#)]
10. Dinarello, C.A. Thermoregulation and the pathogenesis of fever. *Infect. Dis. Clin. N. Am.* **1996**, *10*, 433–449. [[CrossRef](#)]
11. Van Vliet, M.; Donnelly, J.P.; Potting, C.M.J.; Blijlevens, N.M.A. Continuous non-invasive monitoring of the skin temperature of HSCt recipients. *Support Care Cancer* **2010**, *18*, 37–42. [[CrossRef](#)] [[PubMed](#)]
12. Dambrosio, N.; Gannon, J.; Eager, J.; Sylvan, D.; De Lima, M.J.; Levitan, N.; Malek, E. Feasibility of monitoring of body temperature for patients undergoing stem cell transplant or intensive chemotherapy. *J. Clin. Oncol.* **2017**, *35*, e21630. [[CrossRef](#)]
13. Jordán-Núñez, J.; Miró-Martínez, P.; Vargas, B.; Varela-Entrecanales, M.; Cuesta-Frau, D. Statistical models for fever forecasting based on advanced body temperature monitoring. *J. Crit. Care* **2017**, *37*, 136–140. [[CrossRef](#)]
14. Smarr, B.L.; Aschbacher, K.; Fisher, S.M.; Chowdhary, A.; Dilchert, S.; Puldon, K.; Rao, A.; Hecht, F.M.; Mason, A.E. Feasibility of continuous fever monitoring using wearable devices. *Sci. Rep.* **2020**, *10*, 1–11. [[CrossRef](#)]
15. Cuesta-Frau, D.; Varela-Entrecanales, M.; Valor-Perez, R.; Vargas, B. Development of a Novel Scheme for Long-Term Body Temperature Monitoring: A Review of Benefits and Applications. *J. Med. Syst.* **2015**, *39*, 39. [[CrossRef](#)] [[PubMed](#)]
16. Vargas, B.; Cuesta-Frau, D.; Ruiz-Esteban, R.; Cirugeda, E.; Varela, M. What Can Biosignal Entropy Tell Us About Health and Disease? Applications in Some Clinical Fields. *Nonlinear Dyn. Psychol. Life Sci.* **2015**, *19*, 419–436.
17. Churruca, J.; Vigil, L.; Luna, E.; Ruiz-Galiana, J.; Varela, M. The route to diabetes: Loss of complexity in the glycemic profile from health through the metabolic syndrome to type 2 diabetes. *Diabetes Metab. Syndr. Obes.* **2008**, *1*, 3–11.
18. Ogata, H.; Tokuyama, K.; Nagasaka, S.; Tsuchita, T.; Kusaka, I.; Ishibashi, S.; Suzuki, H.; Yamada, N.; Hamano, K.; Kiyono, K.; et al. The lack of long-range negative correlations in glucose dynamics is associated with worse glucose control in patients with diabetes mellitus. *Metabolism* **2012**, *61*, 1041–1050. [[CrossRef](#)]
19. Varela, M.; Churruca, J.; Gonzalez, A.; Martin, A.; Ode, J.; Galdos, P. Temperature curve complexity predicts survival in critically ill patients. *Am. J. Respir. Crit. Care Med.* **2006**, *174*, 290–298. [[CrossRef](#)]
20. Lundelin, K.; Vigil, L.; Bua, S.; Gomez-Mestre, I.; Honrubia, T.; Varela, M. Differences in complexity of glycemic profile in survivors and nonsurvivors in an intensive care unit: A pilot study. *Crit. Care Med.* **2010**, *38*, 849–854. [[CrossRef](#)] [[PubMed](#)]
21. Sandu, A.L.; Staff, R.T.; McNeil, C.J.; Mustafa, N.; Ahearn, T.; Whalley, L.J.; Murray, A.D. Structural brain complexity and cognitive decline in late life—A longitudinal study in the Aberdeen 1936 Birth Cohort. *Neuroimage* **2014**, *100*, 558–563. [[CrossRef](#)] [[PubMed](#)]
22. Costa, M.D.; Goldberger, A.L.; Peng, C.K. Multiscale entropy to distinguish physiologic and synthetic RR time series. *Comput. Cardiol.* **2002**, *29*, 137–140. [[PubMed](#)]
23. Navarro, V.; Martinerie, J.; Le Van Quyen, M.; Clemenceau, S.; Adam, C.; Baulac, M.; Varela, F. Seizure anticipation in human neocortical partial epilepsy. *Brain* **2002**, *125*, 640–655. [[CrossRef](#)] [[PubMed](#)]
24. Cuesta-Frau, D.; Dakappa, P.H.; Mahabala, C.; Gupta, A.R. Fever Time Series Analysis Using Slope Entropy. Application to Early Unobtrusive Differential Diagnosis. *Entropy* **2020**, *22*, 1034. [[CrossRef](#)]
25. Cuesta-Frau, D. Slope Entropy: A New Time Series Complexity Estimator Based on Both Symbolic Patterns and Amplitude Information. *Entropy* **2019**, *21*, 1167. [[CrossRef](#)]
26. Cuesta-Frau, D.; Miró-Martínez, P.; Oltra-Crespo, S.; Molina-Picó, A.; Dakappa, P.H.; Mahabala, C.; Vargas, B.; González, P. Classification of fever patterns using a single extracted entropy feature: A feasibility study based on Sample Entropy. *Math. Biosci. Eng.* **2020**, *17*, 235. [[CrossRef](#)]
27. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [[CrossRef](#)]
28. Dakappa, P.H.; Rao, S.B.; Bhat, G.K.; Mahabala, C. Unique temperature patterns in 24-h continuous tympanic temperature in tuberculosis. *Trop. Doct.* **2019**, *49*, 75–79. [[CrossRef](#)]

29. Dakappa, P.H.; Prasad, K.; Rao, S.B.; Bolumbu, G.; Bhat, G.K.; Mahabala, C. A Predictive Model to Classify Undifferentiated Fever Cases Based on Twenty-Four-Hour Continuous Tympanic Temperature Recording. *J. Healthc. Eng.* **2017**, *2017*, 6. [[CrossRef](#)]
30. Cuesta-Frau, D.; Schneider, J.; Bakštein, E.; Vostatek, P.; Spaniel, F.; Novák, D. Classification of Actigraphy Records from Bipolar Disorder Patients Using Slope Entropy: A Feasibility Study. *Entropy* **2020**, *22*, 1243. [[CrossRef](#)]
31. Li, Y.; Gao, P.; Tang, B.; Yi, Y.; Zhang, J. Double Feature Extraction Method of Ship-Radiated Noise Signal Based on Slope Entropy and Permutation Entropy. *Entropy* **2022**, *24*, 22. [[CrossRef](#)]
32. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)] [[PubMed](#)]
33. Cirugeda-Roldán, E.; Cuesta-Frau, D.; Miró-Martínez, P.; Oltra-Crespo, S.; Vigil-Medina, L.; Varela-Entrecanales, M. A new algorithm for quadratic Sample Entropy optimization for very short biomedical signals: Application to blood pressure records. *Comput. Methods Programs Biomed.* **2014**, *114*, 231–239. [[CrossRef](#)] [[PubMed](#)]
34. Lake, D.E.; Moorman, J.R. Accurate estimation of entropy in very short physiological time series: The problem of atrial fibrillation detection in implanted ventricular devices. *Am. J. Physiol.-Heart Circ. Physiol.* **2011**, *300*, H319–H325. [[CrossRef](#)]
35. Lake, D.E.; Richman, J.S.; Griffin, M.P.; Moorman, J.R. Sample entropy analysis of neonatal heart rate variability. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **2002**, *283*, R789–R797. [[CrossRef](#)]
36. Sokunbi, M.O. Sample entropy reveals high discriminative power between young and elderly adults in short fMRI data sets. *Front. Neuroinform.* **2014**, *8*, 69. [[CrossRef](#)]
37. Li, H.; Peng, C.; Ye, D. A study of sleep staging based on a sample entropy analysis of electroencephalogram. *Bio-Med Mater. Eng.* **2015**, *26*, S1149–S1156. [[CrossRef](#)] [[PubMed](#)]
38. Cuesta-Frau, D.; Varela, M.; Miró-Martínez, P.; Galdós, P.; Abásolo, D.; Hornero, R.; Aboy, M. Predicting survival in critical patients by use of body temperature regularity measurement based on approximate entropy. *Med. Biol. Eng. Comput.* **2007**, *45*, 671–678. [[CrossRef](#)]
39. Lee, G.; Fattinger, S.; Mouthon, A.L.; Noirhomme, Q.; Huber, R. Electroencephalogram approximate entropy influenced by both age and sleep. *Front. Neuroinform.* **2013**, *7*, 33. [[CrossRef](#)] [[PubMed](#)]
40. Mayor, D.; Panday, D.; Kandel, H.K.; Steffert, T.; Banks, D. CEPS: An Open Access MATLAB Graphical User Interface (GUI) for the Analysis of Complexity and Entropy in Physiological Signals. *Entropy* **2021**, *23*, 321. [[CrossRef](#)]
41. Aboy, M.; Cuesta-Frau, D.; Austin, D.; Micó-Tormos, P. Characterization of Sample Entropy in the Context of Biomedical Signal Analysis. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 5942–5945. [[CrossRef](#)]
42. Jiang, Y.; Mao, D.; Xu, Y. A fast algorithm for computing sample entropy. *Adv. Adapt. Data Anal.* **2011**, *3*, 167–186. [[CrossRef](#)]
43. Molina-Picó, A.; Cuesta-Frau, D.; Aboy, M.; Crespo, C.; Miró-Martínez, P.; Oltra-Crespo, S. Comparative Study of Approximate Entropy and Sample Entropy Robustness to Spikes. *Artif. Intell. Med.* **2011**, *53*, 97–106. [[CrossRef](#)] [[PubMed](#)]
44. Yentes, J.M.; Hunt, N.; Schmid, K.K.; Kaipust, J.P.; McGrath, D.; Stergiou, N. The Appropriate Use of Approximate Entropy and Sample Entropy with Short Data Sets. *Ann. Biomed. Eng.* **2013**, *41*, 349–365. [[CrossRef](#)]
45. Lu, S.; Chen, X.; Kanters, J.K.; Solomon, I.C.; Chon, K.H. Automatic Selection of the Threshold Value r for Approximate Entropy. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1966–1972. [[CrossRef](#)] [[PubMed](#)]
46. Pincus, S.M.; Huang, W.M. Approximate entropy: Statistical properties and applications. *Commun. Stat.-Theory Methods* **1992**, *21*, 3061–3077. [[CrossRef](#)]
47. Swift, A.; Heale, R.; Twycross, A. What are sensitivity and specificity? *Evid.-Based Nurs.* **2020**, *23*, 2–4. [[CrossRef](#)] [[PubMed](#)]
48. Song, B.; Zhang, G.; Zhu, W.; Liang, Z. ROC operating point selection for classification of imbalanced data with application to computer-aided polyp detection in CT colonography. *Int. J. Comput. Assist. Radiol. Surg.* **2013**, *9*, 79–89. [[CrossRef](#)]
49. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **2018**, *17*, 168–192. [[CrossRef](#)]
50. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
51. Fay, M.P.; Proschan, M.A. Wilcoxon-Mann-Whitney or t -test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statist. Surv.* **2010**, *4*, 1–39. [[CrossRef](#)]
52. Dash, S.; Tripathy, R.K.; Panda, G.; Pachori, R.B. Automated Recognition of Imagined Commands From EEG Signals Using Multivariate Fast and Adaptive Empirical Mode Decomposition Based Method. *IEEE Sens. Lett.* **2022**, *6*, 1–4. [[CrossRef](#)]
53. Delgado-Bonal, A.; Marshak, A. Approximate Entropy and Sample Entropy: A Comprehensive Tutorial. *Entropy* **2019**, *21*, 541. [[CrossRef](#)] [[PubMed](#)]
54. Skoric, T.; Sarenac, O.; Milovanovic, B.; Japundzic-Zigon, N.; Bajic, D. On Consistency of Cross-Approximate Entropy in Cardiovascular and Artificial Environments. *Complexity* **2017**, *2017*, 8365685. [[CrossRef](#)]
55. Varela, M.; Ruiz-Esteban, R.; Martínez-Nicolas, A.; Cuervo-Arango, J.A.; Barros, C.; Delgado, E.G. ‘Catching the spike and tracking the flow’: Holter-temperature monitoring in patients admitted in a general internal medicine ward. *Int. J. Clin. Pract.* **2011**, *65*, 1283–1288. [[CrossRef](#)]
56. Varela, M.; Calvo, M.; Chana, M.; Gomez-Mestre, I.; Asensio, R.; Galdos, P. Clinical implications of temperature curve complexity in critically ill patients. *Crit Care Med.* **2005**, *33*, 2764–2771. [[CrossRef](#)]