the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

Research

The application of structural entropy in tissue based diagnosis

Gian Kayser[*)], Jürgen Görtler[**)], Cleo - Aron Weis[***)], Stephan Borkenfeld[+)], Klaus Kayser[+++)]

[*)] Institute of Surgical Pathology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany
**) IBM Global Markets, Systems, Frankfurt, Germany

***) Institute of Pathology, University Medical Centre Mannheim, University of Heidelberg, Mannheim, Germany
+) International Academy of Telepathology, Heidelberg, Germany

+++) Institute of Pathology, Charite, Berlin, Germany

Abstract

Background: Entropy belongs to the few basic measurable entities in nature. It measures the distance of a closed or open 'statistical' system from its present stage to its final stage, and analyzes the probability distribution of the included elements, independently from their meaning. The development can be predicted by use of an 'ideal transformation', i.e. additive formula (for example Shannon's entropy) or by mathematical derivatives such as the more generalized q – entropy, for example Tsallis entropy. Herein, the internal structures of the system are described which include so – called macro – systems. They are created by individual elements or basic micro – systems, and transformed into essentials of tissue – based diagnosis.

Entropy and neighborhood: The basic entropy approaches consider a spatially force – independent system, i.e., the calculation of the elements' probability distribution does not take into account the formation of macro – systems, or the position of the individual elements within the system or between individual elements. The receiver of an informative signal cannot distinguish whether it has been generated in the center or at the boundary of the system. Only the signal's probability within the information chain and the formation of the chain are informative.However, neighborhood plays an important role in development, maturation, degradation, and dissolution of biological systems. Most cells are generated by cellular division and neighboring cells are more similar in morphology and function than non-neighboring cells. This observation also holds true for 'higher order' biological systems such as animal colonies, forests, or even human societies. Thus, a potentially successful approach of estimating the development of a biological system should include neighborhood definitions and considerations.

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

Neighborhood conditioned (MST) entropy and entropy flow: Any definition of a neighborhood condition is based upon the distances between different elements, called objects. The distances can be weighted by additional object features, might be 'directed', or might include certain 'shadow' conditions (hidden behind another object). The most frequently used algorithm has been introduced by Voronoi in 1902. It can be successfully formulated in graph theory and derived approaches. In microscopic morphology, the construction of weighted minimum spanning trees (MST) is a convincing approach. Living biological systems are open and not closed. They exchange energy, information, and directives for future development with their environment. They have to stabilize their own entropy level against that of their environment. The mandatory entropy exchange or entropy flow from the individual element into its environment or vice versa reflects to the system's stability and impact on its environment.

Tissue – based diagnosis and entropy: Tissue – based diagnosis includes all technical procedures to ascertain a 'medical diagnosis', such as microscopic, electron microscopic investigations, gene analysis, proteomics, syntactic structure analysis, liquid biopsies, etc. Herein, the transformation and applicability of the entropy approach are described and discussed.

Keywords: entropy, entropy flow, tissue – based diagnosis, prognosis, morphometry, karyometry.

Introduction

Our understanding of disease and health has changed during the last decades. Both items are understood to belong to the same equilibrium stage of a living system that is regulated by internal and external boundary conditions [1-3]. The 'control of feedback systems' remains effective and provides a structural continuity as long as the boundary limits are not overrun. Otherwise it will turn into forward mechanisms and require additional forces to become stabilized again [4, 5].

Derivatives of this theory are, for example, theories of aging, which explain the increasing instability or vulnerability of live with increasing age or theories of maturation and failures at developmental stages [6-8].

In practice, the boundary conditions are measured by clinical investigations. The results are classified and transferred into applicable diagnoses, which serve for locally available and appropriate treatment [9-12]. The diversity of the patients' environment is partly compensated by globalization or 'health tourism' whose entropy analysis might also be a useful directive approach [13, 14].

Thus, each diagnosis includes a 'treatment advice', which is finally digitized (surgery <yes<>no>, antibiotics <yes<> no>, etc.), the aim to prolong live and ascertain certain live conditions such as analgesia, satisfaction, or live quality [4, 15, 16].

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

The meaning of a medical diagnosis has changed too in close association with its theoretical interpretation. Diagnoses derived from tissue examinations tried to correlate morphological characteristics with the outcome and potential useful therapeutic interactions [3, 10, 17].

They are now a days extended by so-called functional characteristics [3, 4, 17-19]. These include macromolecule (protein) expressions and interactions that serve for information transfer and release of forces to alter structures at or below the cellular level [3]. Examples include predictive diagnoses, individualized (targeted) diagnoses and liquid biopsies. Liquid biopsies measure functional characteristics (molecule fractions) of a local disease (solid tumor) in the peripheral blood, and not at its spatial origin.

The investigations to further understand a disease and derive suitable therapeutic actions occur in the molecular / genetic 'world' and drop out of conventional diagnosis. In addition, they promote thoughts that focus on measurement of appropriate characteristics to reliable forecast development and impact of therapeutic efforts. The entropy approach seems to fulfill all mandatory features of such a successfully assessable tool [3, 4, 15, 18, 20].

Herein, we will explain, describe and interpret the potential contribution of the entropy approach in tissue – based diagnosis and derived therapy.

Definition of entropy

One aim of natural sciences is to describe and explain the present stage of an observable system, and to forecast its future development. Man has developed several tools that successfully map specific system properties in practical transformations. These can be distinguished in measurable attributes and algorithms (formulas). For example, useful attributes of a thermodynamic system are its pressure, temperature, volume and (free) energy. Any appropriate algorithm would try to describe the development of the system, its final stage, and to compare the result with our observations [15, 21-25]

The entropy is such a basic and system specific measure, which is based upon statistics of internal characteristics of the system. It can be applied for systems of different physical realizations including linguistics, thermodynamics, acoustics, optics, or quantum physics [13, 26-29].

Frequent realizations include classic mechanics (ideal gas, reversible and non-reversible thermodynamics), computation (numbers, vectors, etc), mathematics (geometry, image analysis), sociology (income, age, etc. of defined populations), or biology (self organization of macromolecules) [4, 23, 26, 30].

In thermodynamics, the entropy S is a state variable and corresponds to the differential of transported heat divided by the absolute temperature

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

$$\delta S = \delta(Q)/T + \delta(Wdiss)/T \qquad (1)$$

δ(Q) = additive heat of a closed, isolated system; **δ(Wdiss)** = inside the system created heat.

The generalized formula is: $S_{q,r} = [1- (\Sigma_i p_i^q)^r]/[r(q-1)]$ **(2)**

A closed system is described by: $\delta(Q)/T \geq 0$; and increases its entropy until it reaches a final stage or the corresponding statistical distribution of its elements [24].

An open system can, in addition, import or export entropy through its surface.

Independently from its nature, any of these systems consists of numerous distinct properties ***i*** each being measured with a probability p_i. Proposing a strong chaos (independent elements) the distance from its end stage can be computed to

$$S = k*\Sigma(p_i * ln\ 1/p_i) \qquad (3)$$

(**K** = constant, so-called Boltzmann constant), with the condition $\Sigma(p_i)=1$.

If the probability ***pi*** is only related to the total number **N** of observable inner (micro) states (for example 6 in case of a usual die), then $p_i = N_i/N$ holds, and we will obtain the so – called Boltzmann – Gibbs distribution of p_i. The entropy ***S*** is maximum if the ***pi*** do not differ, i.e., once the system has reached its final stage.

The entropy concept is based upon 'inside elements' that are subject of statistical measures. Therefore, it does not make sense to apply the entropy concept to an individual object; such as an individual person, animal, organ, vessel, cell, gene, budget, etc. It is a measure of the 'behavior of the total system', or of 'element agglutinations' (macro-stages) inside the system [13, 15, 18, 20, 23, 31, 32, 33].

Examples of applications include procedures to obtain

- the direction of a process in thermodynamic systems (prognosis, described by the direction of a vector);
- the distance of a system from its final stage (time to death, described by scalar of a vector);
- the amount of energy not available for work during a certain process (impact of therapeutic interactions, described by the projection of a vector)

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

- the disorder of a system (concurrent diseases, described by the number of different vectors).

The prerequisite of equation (3) is the assumption that 'no interacting forces' between the statistical measures exist, which is equal to a 'strong chaos', or the so – called Boltzmann – Gibbs distribution. In this case the statistical elements occupy no space, and the entropies of different systems (including macro-stages) can be simply added as shown in formula (4).

$$S\,(A+B) = Sq(A) + Sq(B) \qquad (4)$$

In reality, however, the statistical elements frequently occupy a non negligible space, and the entropy concept becomes more complicated: In this case we are dealing with the so – called 'weak chaos', or additional existence of locally (spatial) dependent or consistently element associated forces. These forces might create internal 'agglutinations' which are called macro-stages [4, 20, 31, 34].

How to compute the entropy of individual macro-stages, and that of the whole system? Obviously, the calculation depends upon the nature of the 'interaction forces', or of the 'weak chaos'. Therefore, we have to define certain assumptions.

The simplest preposition would be an additive weighted factor in formula (3) that assumes an equal contribution and nature of the included entropies. It is shown in formula (5)

$$S\,(A+B) = Sq(A) + Sq(B) + (1-q) * Sq(A)*Sq(B). \qquad (5)$$

The procedure is called ***(1-q)*** interaction. The factor ***q>0*** corresponds to a linear distance dependence, the occupation space of system elements, or a relaxation (response) time.

This is an approximation concept and useful, if we do not know the interactions of the inner compartments of the systems. We assume that in this case each system is still composed of those figures that will result in the 'theoretical maximum' of the (macro) system [25]; or that we can compute the entropy of our (macro) system in relation to the proposed small deviation **q (0<q<1)**.

Herein, the entropy still remains a measure of the system's distance from its final stage, and the maximum q - entropy still reflects the most probable distribution of inner stages. Tsallis et al. demonstrated that the maximum q – entropy describes also the 'behavior' of (macro) systems accurately. It can easily be calculated for different approaches with the help of Naudts logarithm [25, 35].

Individual event and probability

In principle, the entropy concept can be applied to a broad variety of systems with "internal statistical properties"**,** i.e., which contain elements that are subject to statistical computations.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

The nature of the 'statistical elements' is irrelevant. Therefore, entropy is a suitable entity to analyze systems of a broad variety.

Its advantage is also its disadvantage: it does not make sense to estimate the probability of the presence of an individual element, which can only be answered by saying <yes or no>; i.e., it is senseless since we do know it already.

For example, one has just caught a fish or missed it. Nothing is in between. The question 'Have you caught a fish?' cannot be answered by any entropy concept

However, it makes sense to forecast man's future fishing prosperity by probability calculations:

For example, the fish has a high risk to be caught if it swims close to the fishing rod. Thus, it makes sense to associate statistical probabilities and derived terms, such as entropy, to future expressions of individual elements (fish) of the system (shoal). Naturally, it reflects to the development of the whole system and its potential compartments too.

If we associate statistical properties with the system's elements (for example distance of the fishes to the fishing rod, food availability, experience of the fisherman, etc.), and if their distribution can be estimated, the general concept of entropy can be used to forecasting the development of the derived macro-system (number of fishes), and its influence on boundary conditions (disappointment of the fisherman and clear out of the fishing rod).

Statistics of basic populations: macro- and micro- stages

Taking a shoal, a tissue compartment, a forest, or any of so – called macro-systems that are composed of numerous individual events, we might detect that their basic elements are "aggregated" to individual classes too. This distribution characterizes the distinct macro-stage at time of measurement.

Kolmogorov proposed an axiomatic approach of non-overlapping elementary events e_i to describe the statistical properties of the macro system as follows:

Each individual event $\mathbf{<e_i>}$ is characterized by a number $\boldsymbol{p_i}$ within the range $\mathbf{0 \leq p_i \leq 1}$ and fulfils the condition $\boldsymbol{\Sigma\{p_i\} = 1}$ [17]. The number $\boldsymbol{p_i}$ reflects to the probability of $\boldsymbol{e_i}$ within the macro - system (for example fish within the shoal). Often, the elementary events aggregate to similar sub-compartments $\boldsymbol{c_i}$ and form a characteristic class (fishes of certain size, age, gender, species, etc.).

The question arises: What are the stages and the most likely distribution of $\boldsymbol{c_i}$?

Two different methods of measurements do exist, namely a) to undertake multiple serial tests with the same sub-compartment, or b) to measure several identical sub-compartments contemporarily to answer the mentioned question.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

Usually, both strategies will give the same result. For example, we play either with one die several times or with multiple dies only once, and assign the number to each observed event (1 – 6). The probabilities will be the same under ideal conditions. If we play with several *(D)* dies, each die corresponds to a macro-stage. The total number of possible micro-stages results in $N_t = 6^D$, and the total number of possible macro-stages

$$Mt = \binom{6+D-1}{6-1}$$

Each macro - stage W can be calculated according to Boltzmann

$$W = N!/(\prod_{1}^{6}(Ni!)) \qquad (5)$$

Using Sterling's formula $ln(W) = ln(N!) - \Sigma\{Ni*ln(Ni)!\}$, the expression finally results in

$1/Nt*ln(W) = -\Sigma\{pi*ln(pi)\} = S$ (entropy).

In tissue – based diagnosis, such calculations can be used to assign glands, vessels, nerves, T-cell agglutinations, bacillus colonies, etc. with macro-stages, and the individual cells (goblet, endothelial, nerve sheet cells, etc.) with micro–stages [4, 15, 17, 36, 37].

Entropy and tissue structures

In due considerations tissue structures correspond to macro-systems: Their entropy as well as their contribution to the total (host) system entropy can be measured. Their distances in between can also be transformed into a 'distance entropy system', which might include macro-systems (vessels, glands, membranes displaying with / without receptors, etc.) too.

The condition of the total (host) system to be fulfilled remains $\Sigma(p_i) = 1$. It can be extended by additional conditions of an included macro-system to $\Sigma(p_i*a_i) = A\sim$. $A\sim$ (or $N* a\sim$) stand for parameters of the macro-system, such as total energy, total of system information, total of die or cell numbers, etc.

As shown by Voß [31, 38] the approach results in a general formula

$$S = C* ln\ W + Ca \qquad (6)$$

C and Ca are constants which depend on the micro- stages (events) of the host system and their parameters (aim of measurement). For example in linguistics upon vocals, consonants,

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

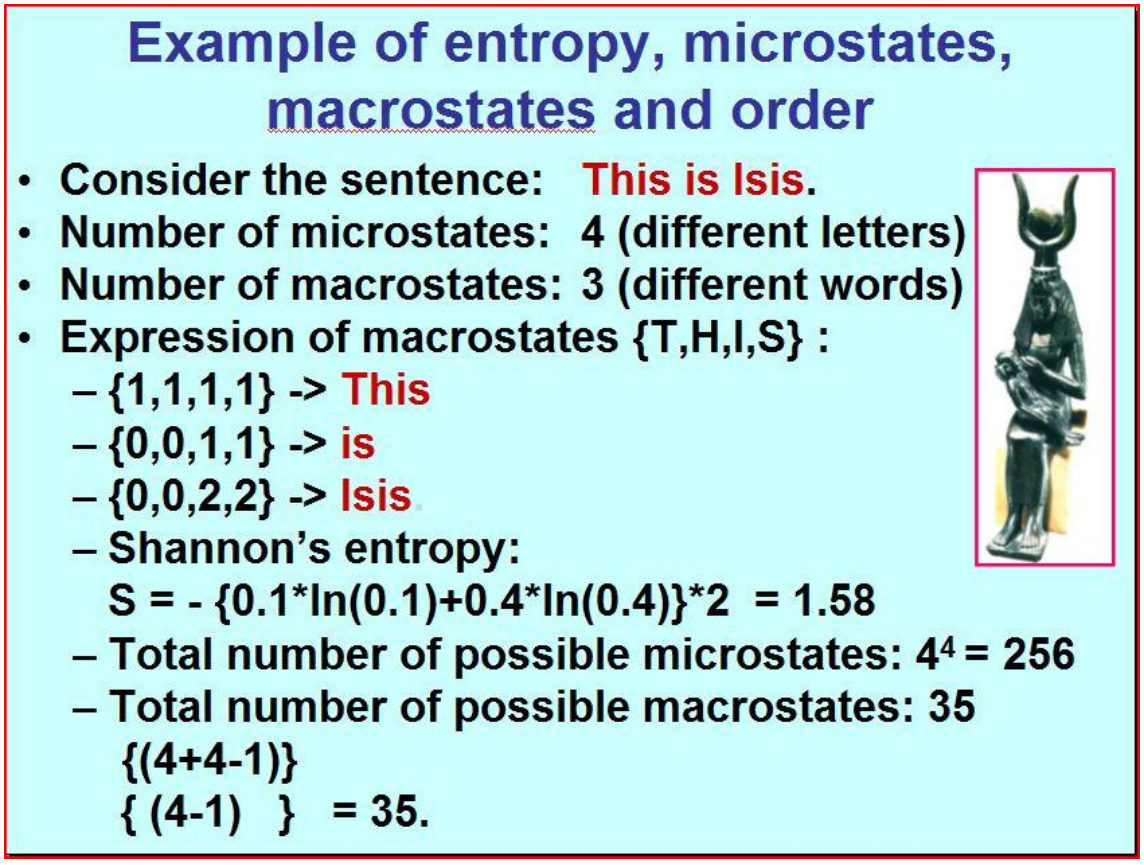
length of words, etc. or, in molecular genetics upon the sequence of nuclear acids or in pathology, upon the expression of receptors in organ structure and function [1, 20, 60, 68-72].

In other words, herein the entropy is a statistical description of a homogeneous macro-stage which is built by micro-stages of the host system (epithelial cells of a gland, endothelial cells of a vessel, etc.) [4, 15, 17, 36, 37].

Two principal entropy concept approaches exist to investigate in tissue characteristics: a) algorithms to measure its entropy of micro-stages and included macro-stages; b) to include 'distance measures' of micro-stages and macro-stages and create a transformed system where the 'distances' serve for statistically accessible events [5, 16, 19, 20, 39].

An example how to measure the entropy of macro-stages and that of the host system without considerations on its structure is shown in <figure 1> [20]:

Figure 1: The sentence <this is isis> includes four letters (micro-stages) and 3 words (macro-stages); The four different statistical elements (letters) can create a maximum of 35 different macro-stages (words); A total of 256 elements (letters) is needed to realize all 35 macro-stages.



*

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

We will reach a different situation, if we take into account distant measures or an included structure which is depicted in <figure 2>, [20]:

Figure 2: The order of the macro-stages (words) and their neighborhood condition play an important role in information and understanding.;The amount of calculated entropies corresponds to this observation.

Entropy, microstages, macrostages and structure

- **Consider the sentence: This is Isis**
- **Number of microstages: 4 (different letters)**
- **Number of macrostages: 3 (different words)**
- **Distance between microstages {0,1}**
- **We get: {t0h0i0s1i0s1i0s0i0s1} or 3 structures.**
- **Shannon's entropy of (This is Isis)**
  Structure 1: 1.38 structure 2: 0.69 structure 3: 1.38
- **Entropy difference between structures:**
  Structure 1: -0.69 -> structure 2: +0.69 -> structure 3
- **Entropy difference of (Is this Isis?)**
  Structure 1: 0.69 -> structure 2: 0 -> structure 3.
- **Total entropy of macrostructures:**
  – **Shannon's S = 3.45; and: $Sm = \Sigma\{p_k * \ln(p_k)\} = 0.22$**

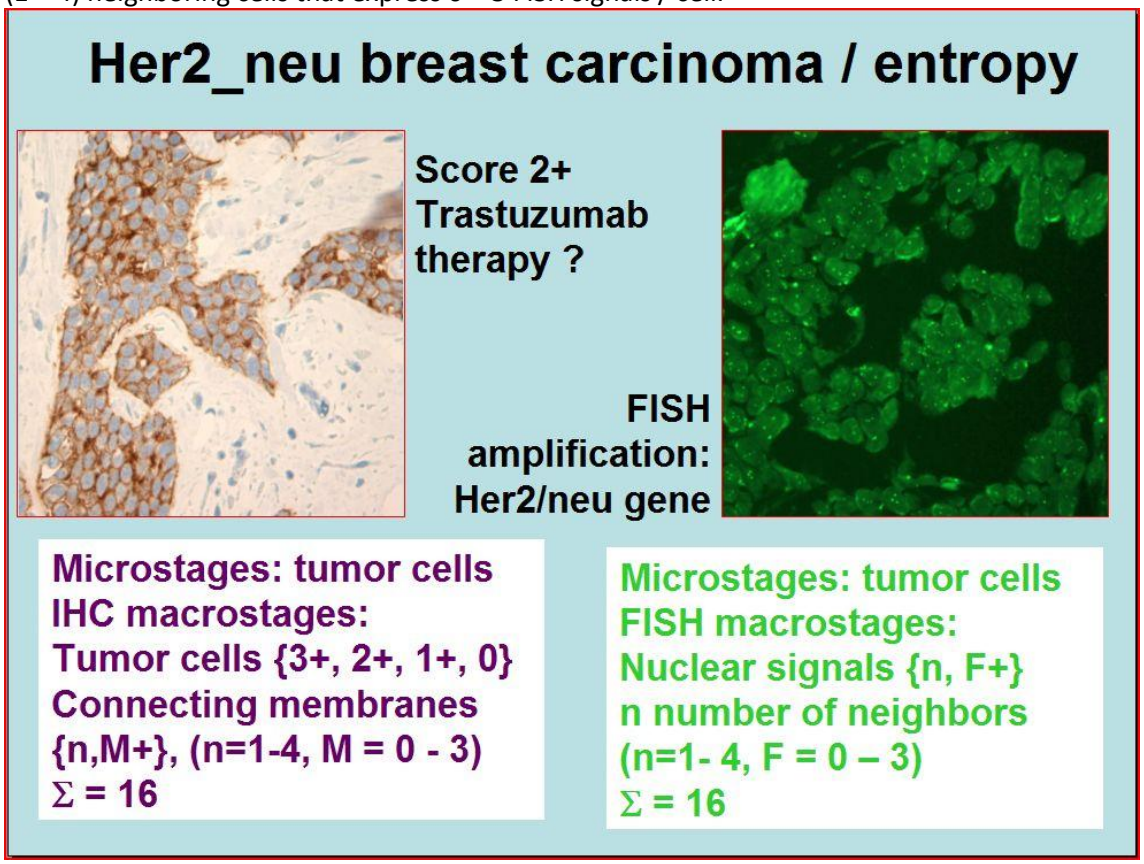
*

The example demonstrates the influence of 'neighborhood' between macro-stages. Assuming strong chaos of the system and its macro-stages, the different entropies can be added, for example from the left (start of the sentence) to the right (end of the sentence). The 'distance entropy of the statement (this is isis) results to -0.69 in contrast to the question (is this isis?), which is equal to zero.

Additional examples are shown in <figure 3, figure 4, and figure 5>. Herein, tumor cells are proposed to be micro-stages that form macro-stages according to the her2neu expressions of their neighbors (IHC grading and FISH). The data display with maximum entropy values which accurately correspond to the visual score 1, or 2 – 2.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

Figure 3: Application of the entropy concept in targeted therapy;
Example of IHC and FISH her2_neu investigation in invasive breast carcinoma (visual score = 2);
Right side: IHC expression scores (0 – 3) correspond to four macro-stages with N = 16 micro–stages (1 – 4) connecting membranes that express IHC stain intensity scores (0-3)
Left side: (FISH signal scores (0 - 3) correspond to four macro–stages with N = 16 micro–stages (1 – 4) neighboring cells that express 0 – 3 FISH signals / cell.



*

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

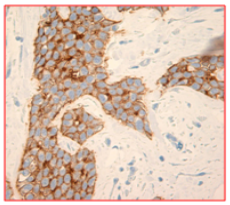
Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

The obtained results of the entropy calculation are depicted in < figure 4> and <figure 5>.

Figure 4: Measures derived from the entropy concept based upon proposed macro-stages (neighboring cells with scored IHC membrane intensity as shown in <figure 3>.

Results: IHC

Microstages: Σ tumor cells: 320
M(0+) = 22, M(1+)=66, M(2+)=214,
M(3+)=38, N = No neighbors
Entropy: 1.03 Macrostages: 16



	N	1	2	3	4	Entropy
M	*0*	9	13	0	0	0.58
M	*1*	28	20	8	10	1.19
M	*2*	16	68	82	47	1.19
M	*3*	19	16	3	0	1.09

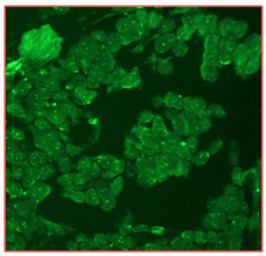
*

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

Figure 5: Measures derived from the entropy concept based upon proposed FISH macro-stages (neighboring cells with scored FISH signals as shown in <figure 3>.

Results: FISH

Microstages: Σ tumor cells: 120
FISH(0+)= 8, FISH(1+)=16,
FISH(2+)=46, FISH(3+)=50
Entropy: 1.18 Macrostages: 16



	N	1	2	3	4	Entropy
FISH	*0*	5	3	0	0	0.35
FISH	*1*	6	4	6	0	1.66
FISH	*2*	6	9	17	14	1.13
FISH	*3*	9	23	13	5	1.23

*

The exemplarily performed calculations use a specific neighborhood condition which includes Voronoi's condition, Dirichlet's tessellation, and weighted IHC edges (<figure 3, figure 4>), or weighted vertices (<figure 3, figure 5>). The computation results in digital data that might serve for automated scores and detailed insight in tumor cell signaling. They can be extended in entropy flow calculation.

Structural (MST) entropy and entropy flow

Considerations on neighborhood play an important role in natural sciences [32, 33, 40-43]. A strong association of similar events, objects, and structures is frequently noticed, and can be assumed to be a general law in nature. It also reflects to information or signal exchange and understanding: the closer elements or events are located the more frequent, easier and accurate are their information exchange and derived actions.

In biology, neighborhood is a prerequisite to create structures, and, vice versa, neighborhood considerations might detect structures.

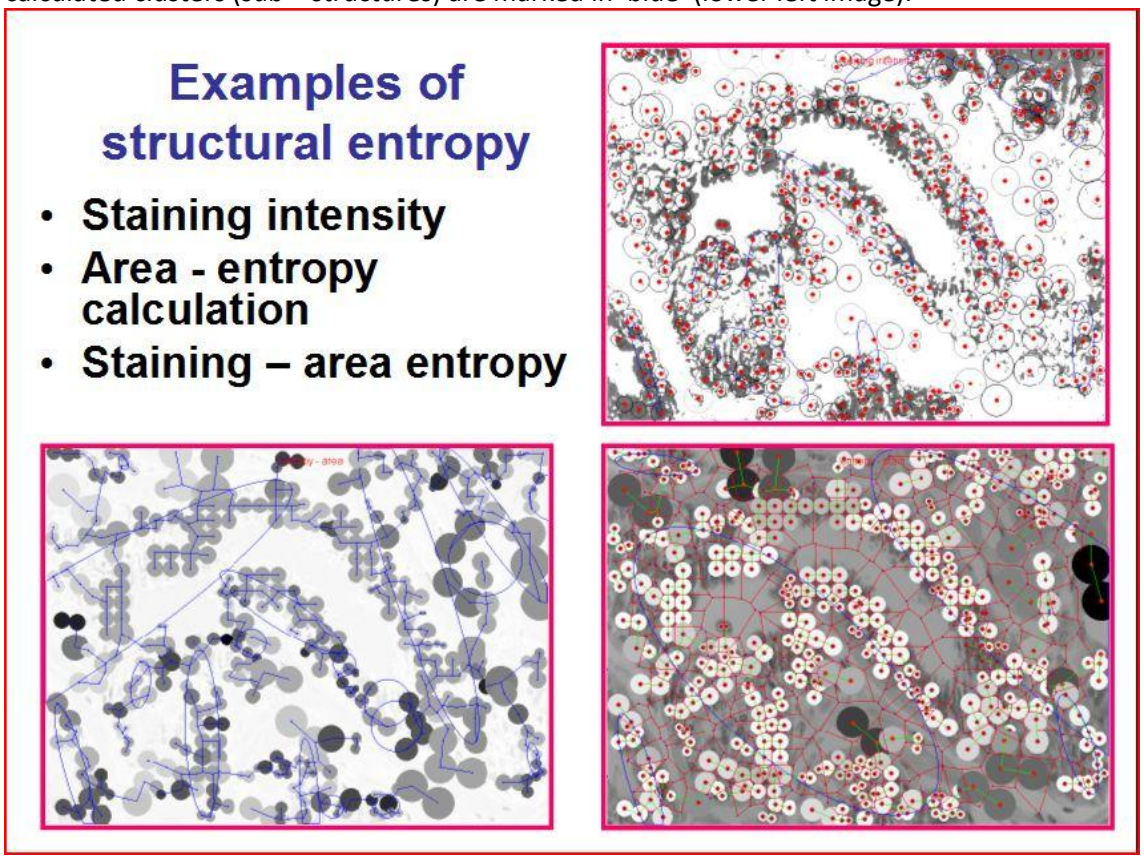
Voronoi's neighborhood condition and its derivative, Dirichlet's tessellation are the most frequently applied techniques in two dimensional spaces [18, 19, 31, 36, 38, 44]. They can be

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

extended by O'Callaghan's condition which takes into account lower and upper distances between proposed neighbors [42, 45]. This approach can simulate tissue spaces of different nature such as air spaces, collagen, vessels, etc. An example is depicted in <figure 6>.

Figure 6: Example of entropy calculations based upon Voronoi's neighborhood condition and O'Callaghan's limitation; The construction of the minimum spanning tree (MST) entropy in combination with two node scalars (size, staining intensity, DAB, CEA) is shown [19, 20, 46]; The calculated clusters (sub – structures) are marked in 'blue' (lower left image).



*

The basic formula of this approach can be written:

$$E(MST) = 1/N * \Sigma[p^*(mst)* ln[p^*(mst)]] \qquad (7)$$

$E(MST)$ = structural entropy, N = number of elements, $p(mst) = SV_{ik} * SE_{ij}$

SV_{ik} = set of parameters associated with the individual neighbouring vertices i, k, i.e.

$$SV_{ik} = [(N-1) * \delta(v_i - v_k) / \Sigma(\delta(v_i - v_k)]^{*2}$$

$$SE_{ij} = [(N-1) * \delta(e_{ij}) / \Sigma(\delta(e_{ij})]^{*2}$$

$\delta(v_i - v_k)$, $\delta(e_{ij})$ = distance of weights between nearest vertices v_i, v_j, or connecting edges e_{ij}.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

The analysis of internal entropy neighborhood details can be expanded to calculate the entropy differences between substructures. Knowing the surface of the substructures, their entropy differences and distances to their neighboring structures we can easily compute the entropy flows which are $\Delta H = \delta W / S * t$. ($\Delta H$ = entropy flow, δW = entropy difference, S = surface compartment of the substructure in relation its neighbor, t = normalized time period (1).

Onsager's theory predicts that the entropy flow tries to reach a minimum, which can be formulated as follows:

$$dH/(S*t) = [\{\delta(Q/T) + \delta(Wdiss)/T\}_{sys} - \{\delta(Q/T) + \delta(Wdiss)/T\}_{env}]/[S*t] \rightarrow minimum; \quad (8)$$

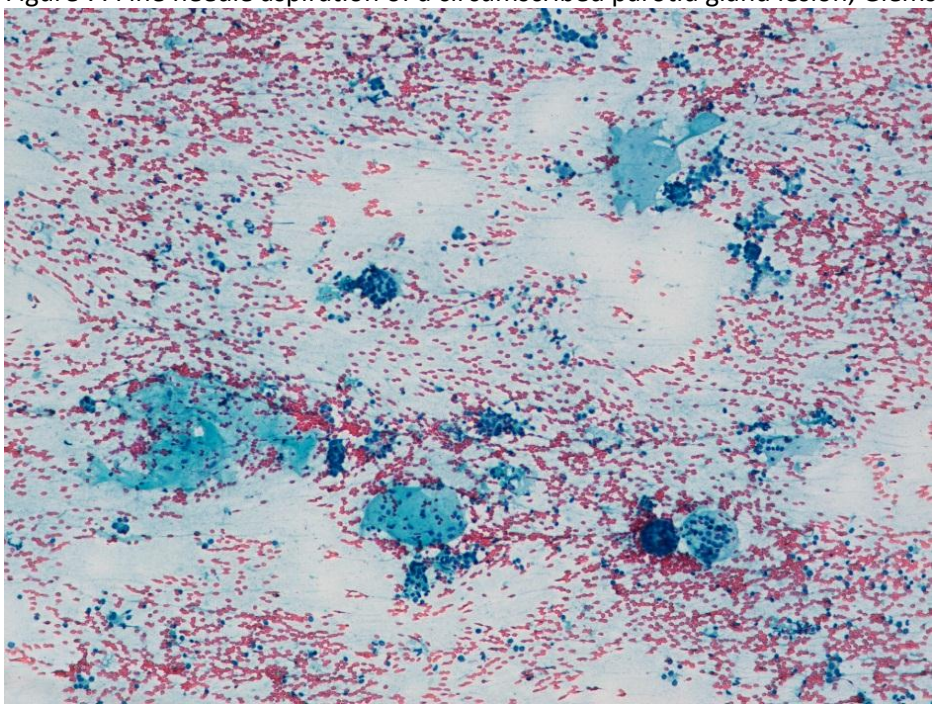
The entropy flow can also be calculated for circumscribed tissue compartments such as solid tumors, lymph nodes, etc. [17, 33, 36, 43, 47]. Kayser et al. reported a close correlation of the entropy flow to the survival of patients with surgically treated lung carcinomas and intra-pulmonary metastases [48-51]

Fine needle aspirations or cytology specimens might be evaluated in a similar manner. An example is shown in <figure 7 and figure 8>. Cellular clusters are separated from the environment by layers of pixels that contain different gray values. The density and width of the boundary shells are statistical elements and thus subject for entropy considerations. Such an approach might be an appropriate method to estimate the glare effect or artifacts [52-54].

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

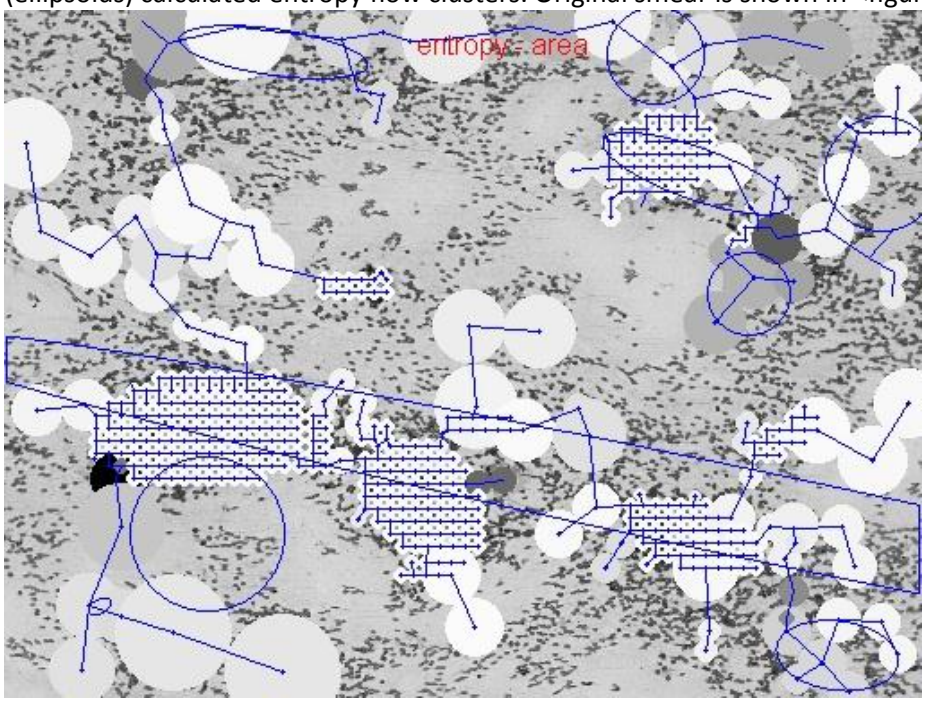
Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

Figure 7: Fine needle aspiration of a circumscribed parotid gland lesion; Giemsa, x 10.



*

Figure 8: Computation of the entropy based upon cell size and neighborhood. Intensity of the gray color indicates the amount of entropy, graphs (in blue) connected aggregates, circles (ellipsoids) calculated entropy flow clusters. Original smear is shown in <figure 7>.



*

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

The entropy concept and image primitives

Pathologists are used to evaluate and diagnose microscopic images. They utilize their knowledge of disease and anatomy and incorporate their experience. They have been trained to recognize certain tissue compartments such as cell types, glands, nerves, vessels, etc.

In other word, the tools to state of a microscopic diagnosis are a 'product' of image content information (ICI) and external (pathologist's) knowledge (PK). Image content information is the information of an image that can be 'extracted' without any external knowledge, i.e., to segment and identify objects without any interpretation what the extracted image objects correspond to [3, 17, 19, 47]. The external information corresponds to the knowledge of the pathologist.

Therefore, the specificity and sensitivity of a pathologist's diagnostic performance is the product of ICI x PK, and the corresponding entropies can be computed

$$S(D) = S(ICI) \times S(PK) \quad (9).$$

How to quantify ***S(ICI)***?

Our considerations on macro-stages and micro-stages might serve to develop appropriate tools. The simplest micro-stages can be defined and classified by gray values of an individual pixel and the simplest macro-stages by one or two dimensional pixel configurations, such as closed and open lines, crosses, or planes such as filled circles, triangles, squares, etc. Such a procedure permits image entropy calculations that are completely independent form any external knowledge [3, 17, 19, 47]. Thus, such an algorithm might analyze and evaluate microscopic images as well as images of forests, cities, wild life or natural disasters.

The entropy approach that addresses the analysis of external knowledge might follow a similar procedure. All known different diagnoses (or distinct expertise events) might be transformed in micro-stages, and the derived actions in macro-stages. The automated entropy calculations will result in estimations how 'severe' the diagnosed disease and 'how potentially successful' the derived actions are. For example, approximately, 860 different pulmonary diseases exist, which induce approximately 20 different interactions [55]. Taking 5 different micro-stages only we will obtain a maximum of 3125 different constellations, and 126 macro-stages. In other word, five image primitives are already sufficient to describe 3125 different diagnoses and 126 derived therapeutic actions, as exemplarily demonstrated in <Figure 1>. Obviously, we have to define additional micro-stages if we want to include more than 126 actions.

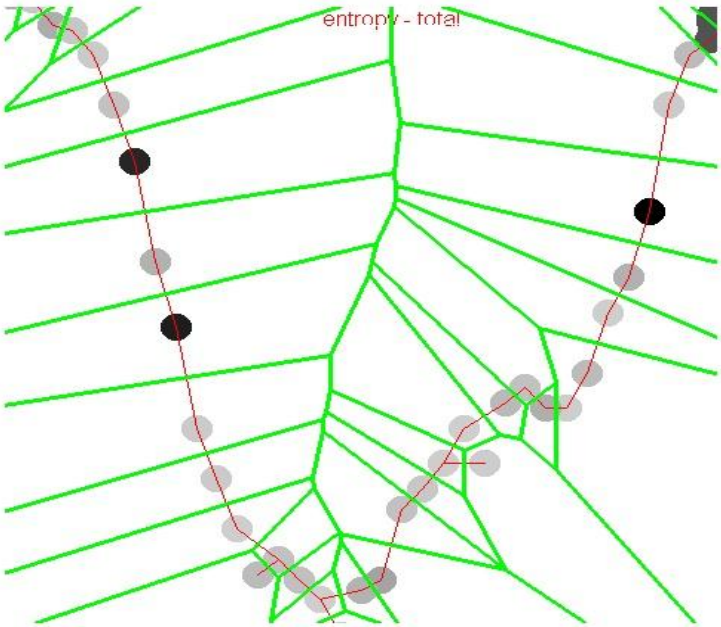
Each pixel gray value might serve for an image primitive, or a predefined width of gray values ($1 < g_i < 10, < 20, < 30, < 40, \ldots$). Consecutively, the entropy approach might be applied to each set

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

of primitives within a gray value range. The next logical step would include the entropy differences between the different sets and the derived entropy flow [56].

Such an entropy approach simulates the thermodynamic forces at 'boundaries' if we assume that strong forces would induce intensive stains (high gray values). This approach permits insight into the thermodynamics of events which cannot or only crudely be repeated. An example is depicted in <figure 9>.

Figure 9: Computed 'entropy flow' in a chicken kidney embryo based upon IHC staining intensities of maturating ducts; (x – axis: gray value (stain intensity); y-axis: calculated entropy difference; color intensity: absolute individual entropy value. The Voronoi cells represent the 'range' of entropy levels.



*

Conclusions

The entropy concept is based upon fundamental characteristics of nature, and therefore unique. It is a tool to analyze statistical events within a broad spectrum of applicable systems and might serve for successful transformation of these events into a principal statistical model. The analysis permits measure and forecast of the system's present stage and its future development.

Applications in tissue – based diagnoses include measures of structures and derived functions within circumscribed tissue lesions, for example tumor cell heterogeneity and its impact on prognosis [3, 17, 56]. Proposed therapeutic actions might be associated as well. It can also be

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

extended to estimate thermodynamic stages within a cellular microenvironment without repetitive investigations.

The entropy concept is of solely statistical nature and not applicable for individual events. Its correct application opens new doors in medical and tissue – based diagnosis. Being a solely system dependent procedure, it can be embedded in appropriate feedback mechanisms and permit fully automated diagnosis and therapy associated algorithms in the near future.

Acknowledgement: The financial support of the Verein zur Förderung des biologisch technologischen Fortschrits in der Medizin is greatfully acknowledged.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

References

1. Kayser K. *Logic and diagnosis.* Methods Inf Med. 1975; **14**(2): 76-80.
2. Kayser K. *Digital Lung Pathology*. Berlin Heidelberg New York: Springer; 1991.
3. Kayser K, Borkenfeld S, Carvalho R, Kayser G. *Structure, Function, and Predictive Diagnosis Algorithms.* Diagnostic Pathology; 2016. **8**.
4. Kayser K, Gabius HJ. *Graph theory and the entropy concept in histochemistry. Theoretical considerations, application in histopathology and the combination with receptor-specific approaches.* Prog Histochem Cytochem; 1997. **32**(2): 1-106.
5. Kayser K, Hoffgen H. *Pattern recognition in histopathology by orders of textures.* Med Inform (Lond). 1984; **9**(1): 55-9.
6. Gabius HJ, Roth J. *An introduction to the sugar code.* Histochem Cell Biol. 2017; **147**(2): 111-117.
7. Kaltner H, Toegel S, Caballero GG, Manning JC, Ledeen RW, Gabius HJ. *Galectins- their network and roles in immunity tumor growth control.* Histochem Cell Biol. 2017; **147**(2): 239-256.
8. Toegel S, Bieder D, Andre S, Altmann F, Walzer SM, Kaltner H, Hofstaetter JG, Windhager R, Gabius HJ. *Glycophenotyping of osteoarthritic cartilage and chondrocytes by RT-qPCR, mass spectrometry, histochemistry with plant/human lectins and lectin localization with a glycoprotein.* Arthritis Res Ther. 2013; **15**(5): 147.
9. Evans P, Wolf B. *Collaboration rules.* Harv Bus Rev. 2005; **83**(7): 96-104.
10. Kayser K. *Quantification of virtual slides: Approaches to analysis of content-based image information.* J Pathol Inform. 2011; **2**: 2.
11. Kayser K, Borkenfeld S, Goldmann T, Kayser G. *To be at the right place at the right time.* Diagn Pathol. 2011; **6**: 2-9.
12. Kayser K, Görtler J, Giesel F, Kayser G. *How to implement grid technology in tissue-based diagnosis: diagnostic surgical pathology.* Expert Opin Med Diagn. 2008; **2**(3): 323-337.
13. Anza F, Vedral V. *Information-theoretic equilibrium and observable thermalization.* Sci Rep. 2017; **7**: 44066.
14. Bailey K. *Sociocybernetics and social entropy theory.* Kybernetics. 2006; **35**: 375-384.
15. Kayser K, Gabius HJ. *The application of thermodynamic principles to histochemical and morphometric tissue research: principles and practical outline with focus on the glycosciences.* Cell Tissue Res. 1999; **296**(3): 443-55.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

16. Kayser K, Hufnagl P, Kayser G, Zink S. *Stratified sampling: Basic ideas and application in pathology.* Elec J Pathol Histol. 1999; **5**: 994-06.
17. Kayser K, Borkenfeld S, Carvalho R, Djenouni A, Kayser G. *How to analyze Structure and Function in Tissue – based Diagnosis?.* Diagnostic Pathology. 2016; **2**.
18. Kayser K, Kayser G, Hufnagl P. *The concept of structural entropy: basic idea and applications.* Elec J Pathol Histol. 2002; **8**: 024-06.
19. Kayser K, Hoshang SA, Metze K, Goldmann T, Vollmer E, Radziszowski D, Kosjerina Z, Mireskandari M, Kayser G. *Texture- and object-related automated information analysis in histological still images of various organs.* Anal Quant Cytol Histol. 2008; **30**(6): 323-35.
20. Kayser K, Kayser G, Metze K. *The concept of structural entropy in tissue based diagnosis.* AQCH, 2007; 29(5): 296-308.
21. Pincus SM. *Approximate entropy as a measure of system complexity.* Proc Natl Acad Sci U S A. 1991; **88(6)**: 2297-2301.
22. Shannon C. *A mathematical theory of communication.* The Bell System Technical Journal. 1948; **27(3)**: 379-423.
23. Eigen M. *Selforganization of matter and the evolution of biological macromolecules.* Naturwissenschaften. 1971; **58**: 465-523.
24. Gibbs J. *Elementary Principles in statistical Mechanics developed with especial reference to the rational. Foundation of Thermodynamics.* New York: Yale University Press; 1902.
25. Tsallis C. *Entropic nonextensivity: a possible measure of complexity.* Chaos, Solitons and Fractals. 2002; **13:** 371-391.
26. Gunduz G, Gunduz U. *The mathematical analysis of the structure of some songs.* Physica A. 2005; **357**: 565-592.
27. Zhang Y, Yang Z, Li W. *Analyses of urban ecosystem based on information entropy.* Ecological Modelling. 2006; **197**: 1-12.
28. Fradkov AL. *Horizons of cybernetical physics.* Philos Trans A Math Phys Eng Sci. 2006; **375:**2088.
29. Souza AM, Nobre FD. *Thermodynamic framework for the ground state of a simple quantum system.* Phys Rev E. 2017; **95**(1-1): 012111.
30. Beretta GP. *Steepest entropy ascent model for far-nonequilibrium thermodynamics: unified implementation of the maximum entropy production principle.* Phys Rev E Stat Nonlin Soft Matter Phys. 2014; **90**(4): 042113.
31. Voß K. *Entropie als statistisches Strukturmaß.* Wiss Z Techn Univs Dresden. 1970; **19**: 1415-1419.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

32. Krige DG, Rand DS. *Some basic considerations in the application of geostatistics to the valuation of ore in South African gold mines.* J South Afr Inst Min Metal. 1976; **77**: 383-391.

33. van Diest PJ, Kayser K, Meijer GA, Baak JP. *Syntactic structure analysis.* Pathologica, 1995; **87**(3): 255-62.

34. Diosi L, Feldmann T, Kosloff R. *On the exact identity between thermodynamic and informatic entropies in a unitary model of friction.* Int J Quantum Information. 2006; **4**: 99-104.

35. Naudts J. *Generalized thermostatistics based on deformed exponential and logarithmic functions.* Physica A,.2004; **340**: 32-40.

36. Kayser K, Molnar B, Weinstein R. *Virtual Microscopy: Fundamentals, Applications, Perspectives of Electronic Tissue-based Diagnosis*. Berlin: VSV Interdisciplinary Medical Publishing; 2006.

37. Kayser K, Stute H, Bubenzer J, Paul J. , *Combined morphometrical and syntactic structure analysis as tools for histomorphological insight into human lung carcinoma growth.* Anal Cell Pathol. 1990; **2**(3): 167-78.

38. Voß K, Süße H. *Praktische Bildverarbeitung*. München, Wien: Carl Hanser Verlag; 1991.

39. Kayser K, Stute H, Tacke M. *Minimum spanning tree, integrated optical density and lymph node metastasis in bronchial carcinoma.* Anal Cell Pathol. 1993; **5**(4): 225-34.

40. Hu X, Wu DD. *Data mining and predictive modeling of biomolecular network from biomedical literature databases.* IEEE/ACM Trans Comput Biol Bioinform. 2007; **4**(2): 251-63.

41. Springelkamp H, Mishra A, Hysi PG, Gharahkhani P, Höhn R, Khor CC, Cooke Bailey JN, Luo X, Ramdas WD, Vithana E, Koh V, Yazar S, Xu L, Forward H,Kearns LS, Amin N, Iglesias AI, Sim KS, van Leeuwen EM, Demirkan A, van der Lee S, Loon SC, Rivadeneira F, Nag A, Sanfilippo PG, Schillert A, de Jong PT, Oostra BA, Uitterlinden AG, Hofman A; NEIGHBORHOOD Consortium, Zhou T, Burdon KP, Spector TD, Lackner KJ, Saw SM, Vingerling JR, Teo YY,Pasquale LR, Wolfs RC, Lemij HG, Tai ES, Jonas JB, Cheng CY, Aung T, Jansonius NM, Klaver CC, Craig JE, Young TL, Haines JL, MacGregor S, Mackey DA, Pfeiffer N, Wong TY, Wiggs JL, Hewitt AW, van Duijn CM, Hammond CJ. *Meta-analysis of genome-wide association studies identifies novel loci that influence cupping and the glaucomatous process.* Nat Commun. 2015; **5**: 4883.

42. O'Callaghan JF. *An alternative definition for neighborhood of a point.* IEEE Trans Comput. 1975; **24**: 1121-1125.

43. Zahn CT. *Graph-theoretical methods for detecting and describing Gestalt clusters.* IEEE Trans Comput. 1971; **20**: 68-86.

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251

44. Prewitt JMS, Wu SC. *An application of pattern recognition to epithelial tissues*. *Computer Applications in Medical Care*. Int: IEEE Computer Society; 1978.

45. Voronoi G. *Nouvelles applications des parametres continus a la theorie des formes quadratiques, deuxieme memoire: recherches sur les paralleloedres primitifs.* J Reine Angew Math. 1902; **134**: 188-287.

46. Kayser K, Radziszowski D, Bzdyl P, Sommer R, Kayser G. *Digitized pathology: theory and experiences in automated tissue-based virtual diagnosis.* Rom J Morphol Embryol. 2006; **47**(1): 21-8.

47. Kayser G, Kayser K. *Quantitative pathology in virtual microscopy: history, applications, perspectives.* Acta Histochem. 2013; **115**(6): 527-32.

48. Kayser G, Baumhakel JD, Szoke T, Trojan I, Riede U, Werner M, Kayser K. *Vascular diffusion density and survival of patients with primary lung carcinomas.* Virchows Arch. 2003; 442(5): 462-7.

49. Kayser K, Kayser C, Rahn W, Bovin NV, Gabius HJ. *Carcinoid tumors of the lung: immuno- and ligandohistochemistry, analysis of integrated optical density, syntactic structure analysis, clinical data, and prognosis of patients treated surgically.* J Surg Oncol. 1996; **63**(2): 99-106.

50. Kayser K, Zink S, Andre S, Schuring MP, Hecker E, Klar E, Bovin NV, Kaltner H, Gabius HJ. *Primary colorectal carcinomas and their intrapulmonary metastases: clinical, glyco-, immuno- and lectin histochemical, nuclear and syntactic structure analysis with emphasis on correlation with period of occurrence of metastases and survival.* APMIS. 2002; **110**(6): 435-46.

51. Weyn B, Tjalma W, Van De Wouwer G, Van Daele A, Scheunders P, Jacob W, Van Marck E. *Validation of nuclear texture, density, morphometry and tissue syntactic structure analysis as prognosticators of cervical carcinoma.* Anal Quant Cytol Histol. 2000; **22**(5): 373-82.

52. Becker RL Jr. *Standardization and quality control of quantitative microscopy in pathology.* J Cell Biochem Suppl. 1993; **17**: 199-204.

53. Haroske G, Dimmer V, Meyer W, Kunze KD. *DNA histogram interpretation based on statistical approaches.* Anal Cell Pathol. 1997; **15**(3): 157-73.

54. Haroske G, Meyer W, Theissig F, Schubert K, Kunze KD. *Remote quantitation server for quality assurance in DNA ploidy analysis.* Anal Quant Cytol Histol. 1998; **20**(4): 302-12.

55. Kayser K. *Analytical Lung Pathology*. Heidelberg, New York: Springer; 1992.

56. Görtler J, Kayser K, Borkenfeld S, Carvalho R, Kayser G. *Cognitive Algorithms and digitized Tissue – based Diagnosis.* Diagnostic Pathology. 2017; 3(1): 248.

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Gian Kayser, Jürgen Görtler, Cleo - Aron Weis, Stephan Borkenfeld, Klaus Kayser; diagnostic pathology 2017, 3:251
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2017-3:251