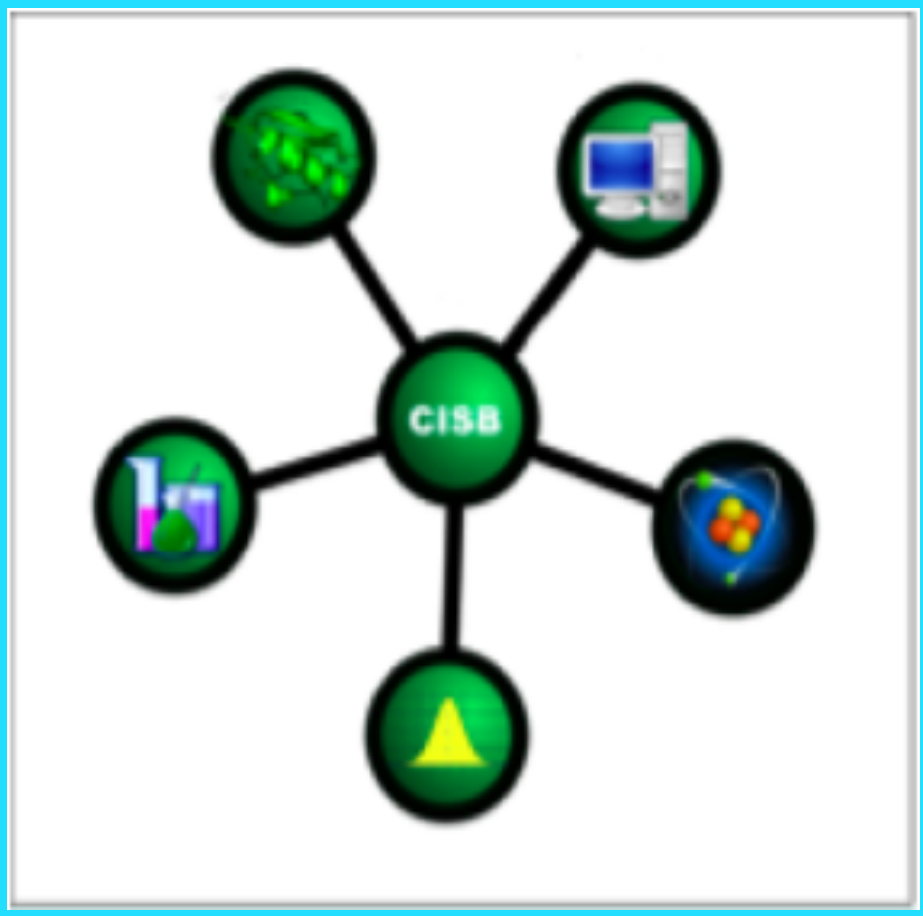
The C.I.S.B. scientific activity: recent and seminal achievements.

Proceedings of an International Symposium - Rome, May 29-30, 2014



Editors: Mauro Colafranceschi, Paolo Sirabella and Alfredo Colosimo

The CISB scientific activity: recent and seminal achievements

Editors: Mauro Colafranceschi, Paolo Sirabella and Alfredo Colosimo - Sapienza University

Rome, May 29-30, 2014

THE CISB SCIENTIFIC ACTIVITY: RECENT AND SEMINAL ACHIEVEMENTS

.

Proceedings of a Symposium (Rome, May 29-30, 2014)

———————-

CISB location downtown (Corso V. Emanuele) [right] and in Palazzo Baleani (floor 1) [left].
The CISB Centre was a direct continuation of a former *Centre of Bioengineering* (CIB), founded in the Seventies of the last century by one of the most inspired rectors of "La Sapienza" University, Antonio Ruberti [1927-2000]. His idea was to provide a *meeting point* to colleagues from biomedical and engineering departments eager to take advantage of complementary theoretical backgrounds and empirical skills in pursuing common scientific targets.

FOREWORD

The *Interdepartmental Centre for the study of Biomedical Systems* (CISB) of "La Sapienza " concluded its life cycle two years ago, after about 25 years of activity. In spite of the limitation concerning any kind of heavy experimental set-up imposed by the location within Palazzo Baleani, a XVII Century building, throughout its whole existence the CISB acted as a "neutral" territory where a wide range of multidisciplinary initiatives could be carried out. Such initiatives included the publication of periodical reports, organization of international meetings, fund-raising, third-level (master, PhD) teaching courses, etc., in the general satisfaction - we believe - of the involved people.

If the 10 years bridging the present and the previous century appeared quite favorable for the interdisciplinary activities carried out at CISB, we easily realize that at present the wind blowing is very different, given the financial restrictions imposed to basic research in favor of the rapid payoff offered by highly specific projects.

However, we were curious to get updated about the evolution of some research lines developed in the last years of the CISB activity, and decided to organize the 2-days Symposium of last May. Thus, colleagues and friends from a number of institutions in Rome and abroad were asked to provide a contribution related to their past connection with the Centre. The result was a significant test of the fruitful collaboration among researchers of my generation catalyzed by the CISB activities. Quite importantly, such activities also provided an excellent environment for full maturation as open-minded and farsighted scientists to PhD students. A good evidence for that is given by the posters presented to the Symposium, also included in these Proceedings, dealing with the

recent work of young colleagues coming from the PhD course in Biophysics of La Sapienza, for many years affiliated to the CISB Centre.

As a final remark, we are inclined to consider the multifaceted typographical style of the volume an acceptable feature more than an annoying flaw: according to the most genuine spirit of the CISB, in fact, we asked the colleagues who generously contributed to the volume to concentrate on the general readability of their papers more than on the obedience to strict formal requirements.

Rome, December 15^{th} 2014

Alfredo Colosimo

———

The "Aula Magna" in Palazzo Baleani

The Symposium was held here.
The somehow austere furnishing of the "Aula Magna", are strikingly contrasted by the exquisite (and fragile) decoration of the roof and the splendid facade of the "Chiesa Nuova" (XVII Century), just on the other side of the road, one can relish from the windows.

General Index

Oral Session

Clavis Universalis: the statistical analysis of shapes

Alessandro Giuliani

Environment and Health Department, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161, Roma, Italy

e-mail: alessandro.giuliani@iss.it

Abstract

A common feature of biology is the presence of a relatively few number of 'privileged forms' out of a huge amount of possible configurations. This happens at different layers of biological organization going from protein three-dimensional structures to gene expression patterns, metabolism, body plans and heartbeat dynamics. The most straightforward definition of form (shape) is: 'an invariant pattern of correlations among a set of landmarks' (Heckman, 1990). Thus, the quantitative description of shapes implies a privileged role played by the 'distance-between-landmarks' computation. The methods based on distance spaces allow for a natural, mainly data driven, approach not requiring any unjustified theoretical hypothesis on the system at hand.

Different methods like multidimensional statistics, complex network analysis, graph theory, converge into the same basic goal of describing shapes in terms of correlation patterns and allow for a natural multiscale appreciation of the studied systems.

The inherent mathematical simplicity of these powerful approaches is stressed in this work.

Keywords: complex networks, principal component analysis, correlation, distances.

1. INTRODUCTION

1.1 Shapes are everywhere

A common, and often misunderstood, feature of living entities is the pervasive presence of few 'privileged forms' getting rid of an apparently bewildering multiplicity of possible configurations in both time and space. Just to name a few:

1) Around one thousand folds are sufficient to categorize the 3D structures of millions of proteins (Govindarajan et al., 1999).
2) No more than 250 kinds of tissues, each corresponding to a strictly invariant gene expression profile, are present in all the metazoans. These profiles arise out of the transfinite number of potential combinations arising from approximately 30000 genes whose expression level can vary across four orders of magnitude (Censi et al., 2010)
3) Less than a dozen of major body plans are sufficient to get rid of the morphology of animals (Hervè et al., 2009).
4) Four main rhythmic activities explain the complex dynamics of heartbeat (Berger et al., 1986).

All the above examples refer to 'shape' as a finite set of relations among elements constraining the number of 'allowed states' of the system at hand.

The collapse of a huge number of theoretically allowed configurations into few effectively observed patterns is the image in light of a rugged energy landscapes (Kauffmann and Weinberger, 1989) in which the observed shapes are the attractor states. This implies the presence of some energy field in which the objects accommodate (Kauffmann and Weinberger, 1989). Different science fields alternatively focus on the field or on the shapes. Chemistry is mainly 'shape oriented': chemical thought, since its pre-scientific alchemic origins, focused on the mutual relations between different entities expressed in terms of natural numbers. This is the case of the concept of valence giving rise to the most widespread topological object of the whole science: the structural formula having atomic species as nodes and covalent bonds as edges (Di Paola et al., 2013). On the other hand, physicists tend to consider structures and shapes as pure consequences of a force field shaping a continuous space, so that the emerging structures are simply "energetically allowed" configurations in the field, with no specific autonomous relevance for the explanation of the studied phenomena (Di Paola et al., 2013).

In the following, I will introduce the basic methodology for the quantitative description of shapes stressing the unique feature of the 'shape-oriented' approach to offer a data-driven, naturally multi-scale, perspective of utmost importance in the analysis of biological systems.

1.2 The definition of shape

Shape (or almost equivalently form or pattern) is one of the most fundamental (and consequently most elusive) concepts in science. The intuitive, while geometrically rigorous, definition of shape deals with the fulfillment of certain constraints linking the different dimensions of a given entity (Di Paola et al., 2012). Thus, a circular shape corresponds to the fulfillment, by a set of points, of the constraint of an invariant distance from a special point called the center, while a triangular shape corresponds to a 180° sum of the internal angles formed by a set of three incident segments. The constraint generates a specific set of correlations among the landmarks of a given shape and consequently a quantitative way to appreciate shape similarities. Figure 1 reports a set of 'similar triangles': the three triangles have identical shape while differing in size, the identity of shape is mirrored by the invariance of the ratios between correspondent edges, i.e by the invariance between the distances between correspondent vertices.

Fig.1

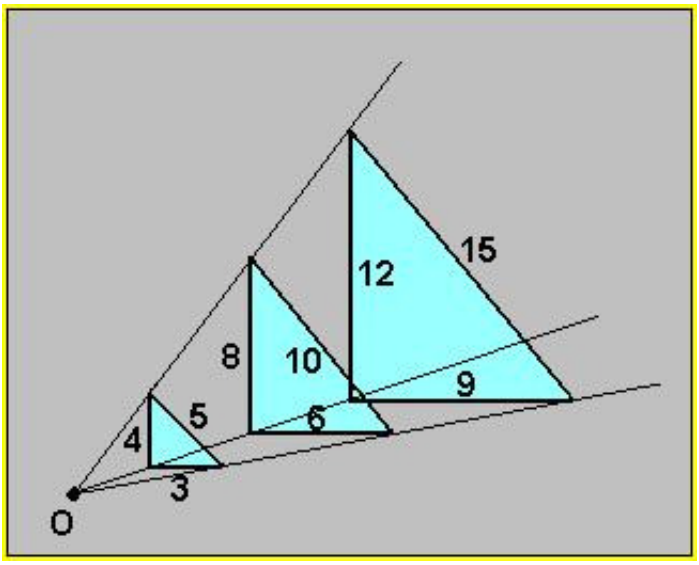


Fig.1: The triangles share the same relations between their edges: 3/5 = 6/10 = 9/15. They are identical in shape and different as for size.

The above line of reasoning is applicable to any kind of complex object by substituting the concept of invariant relation between landmarks with the equivalent concept of between variables correlations. A strictly invariant relation will corresponds to a Pearson correlation coefficient equal to unity, a not-so invariant relation corresponds to a correlation coefficient slightly lower than unity, so opening the way to the quantification of shape modifications (Jolicoeur and Mosimann, 1960).
Carole Heckman (Heckman C., 1990) explicitly states "Earlier work suggested that certain restrictions apply to the geometric configuration that can be assumed by cells. Such restrictions were indicated by the finding that the values of shape descriptors were highly correlated with one another. This was surprising because these descriptors appeared to measure dissimilar geometrical properties of the cell. The present research confirms that the high levels of correlation are due to geometrical constraints on cell shape". It is easy to recognize in these words the definition of shape similarity as a set of correlations: there is no general reason for the existence of a specific correlation between the perimeter and the area of a set of objects, if we observe such a correlation, this implies the objects share the same shape. As a matter of fact any specific shape as a triangle or a circle has a peculiar relation between perimeter and area, going to correlations we can apply the same paradigm to complex objects like proteins or cells (like in the case of Heckman) given we are able to identify suitable descriptors.

2 RESULTS

2.2 Global quantification of size and shape

The above concepts are sufficient to show how to quantify complex geometries in terms of distances. I will start from an apparently purely theoretical example described in (Giuliani et al., 1998) to show this case has sufficient generality to explain relevant biological processes such as gene expression dynamics in tissues.

The idea is to face a complex shape detection problem starting by an explicitly wrong scale. The problem is the reconstruction of the reciprocal positions of European capital cities from their mutual distances from the Latium (the region around Roma) towns. The shape to be described is surely complex: the landmarks, correspondent to the capital cities, have specific relations between them (their mutual distances) but these relations do not obey to a single symmetry rule. The scale is surely wrong (and thus shape detection harder) given Latium is much smaller than Europe.

The distances of each European city from five Latium towns (Roma, Latina, Frosinone, Viterbo, Rieti) were manually estimated by a ruler on a map with a scale 1: ($3*10^6$) in which one mm was correspondent to 3 Kilometers. The much larger extension of Europe ($10.18 * 10^6$ Km^2) with respect to Latium ($17.2 * 10^3$ Km^2) makes the five descriptors (distances from Latium towns) having approximately the same values for each European city. The 'meaningful signal' is embedded into the very small differences of the distances of each European capital city from the various Latium towns. This corresponds to the small deviations from unit of the correlations between the distances from Latium towns of the 33 European capital cities.

This state of affairs is reported in Table1 showing the symmetric correlation matrix among the five descriptors of the system, i.e. the distances from the five Latium towns of the European cities. Table 1 reports the Pearson correlation coefficients between the 33 component vectors correspondent to the distances (measured with a ruler on a 1/3000000 map) between each capital city and the corresponding Latium town.

Table 1

Towns	Roma	Latina	Frosinone	Viterbo	Rieti
Roma	1.0	0.996	0.998	0.998	0.997
Latina		1.0	0.999	0.996	0.988
Frosinone			1.0	0.996	0.992
Viterbo				1.0	0.997
Rieti					1.0

It is important to pay attention to a feature implicit in distance formalization: two cities having the same distance from Latium could nevertheless being very far part from each other simply because they are at the opposite sides with respect to Latium (e.g. Athens and Barcelona, Madrid and Warsaw) so imposing a strong non-linearity to the problem. That is to say that the 'size' component, correspondent to the first eigenvector of Table 1 correlation matrix, will be devoid of any useful shape information, simply reporting the average distance of each city from Latium. Given the small size of Latium this component is expected to be the by far most important as evident by the near to unity correlations between the five descriptors.

Principal Component Analysis (PCA) generates the a five component complete model reported in Table 2 in terms of both component loadings (correlation coefficients between initial values and extracted components) and percentage of explained variance by each component.

Table 2

Towns/Components	PC1	PC2	PC3	PC4	PC5
Roma	0.9997	-0.0137	-0.0184	-0.012	0.0001
Frosinone	0.9973	0.0715	0.0132	0.001	0.0029
Latina	0.9987	0.042	-0.0272	0.0058	-0.0024
Rieti	0.9909	-0.00162	0.0393	-0.0009	-0.0023
Viterbo	0.9964	-0.0837	-0.007	0.006	0.0017
% expl.var.	99.6	0.29	0.057	0.0043	0.0005

As expected, there is a leading principal component (PC1) explaining the almost totality of information (99.6%). This is a size component (Jolicoeur and Mosimann, 1960) having all the variables positively correlated near to unity. PC1 corresponds to the distance of each European city from the centre of Latium (approx. the distance from Roma). The second and third components (PC2,PC3), while explaining a very minor part of information, nevertheless have an interesting loading pattern: the loadings go from Viterbo (most negative) to Frosinone (most positive) while Roma has an intermediate position. This pattern resembles the north-west / south-east spread of Latium towns. Analogously PC3 goes from Latina (most negative) to Rieti (most positive) so resembling west-east ordering. It is worth noting PC2 accounts for a far greater portion of information (explained variance) with respect to PC3, consistently with the fact the main elongation axis of Latium is oriented along north-west / south-east bearing. While PC1 accounts for size information (how much a city is distant from Latium irrespective of its position), PC2 and PC3 are 'shape' components: the different signs of the loadings correspond to the fact that moving along a shape component corresponds to go further from one landmark (town) while approaching another.

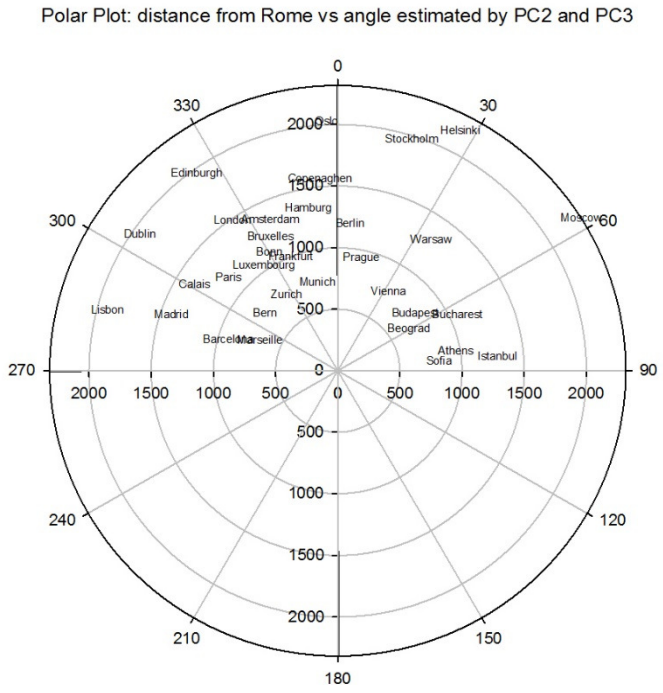
PC4 and PC5 account for a much lower proportion of explained variance and do not show any immediate relation with the actual towns position so they can be interpreted as measurement noise. In (Giuliani et al., 1998) we formally discriminated the 'signal' from the 'noisy' character of PC2,PC3 vs. PC4,PC5. The readers interested in this point can refer to the original paper, here I will concentrate on the signal recognition, i.e. on the above sketched problem of the Europe cities disposition.

We computed, by means of an independent algorithm [PC GLOBE 5.0 , PC GLOBE Inc.], the angle which is formed by the European cities with respect to the north-south axis passing through Roma. We were able to accurately predict this angle by a multiple linear regression having PC2 and PC3 as independent variables, according to the model:

Angle = -4.8306 + 41.097 (PC2) – 27.086 (PC3); r = 0.97, p < 0.0001 (1)

PCA was able to detect the very small signal present in the data relative to the locations of European cities. This is still more clear if we plot the cities disposition on a polar coordinates reference having as radius the distance from Roma (PC1) and as angle the estimate by equation (1) based on PC2,PC3 scores (Fig.2).

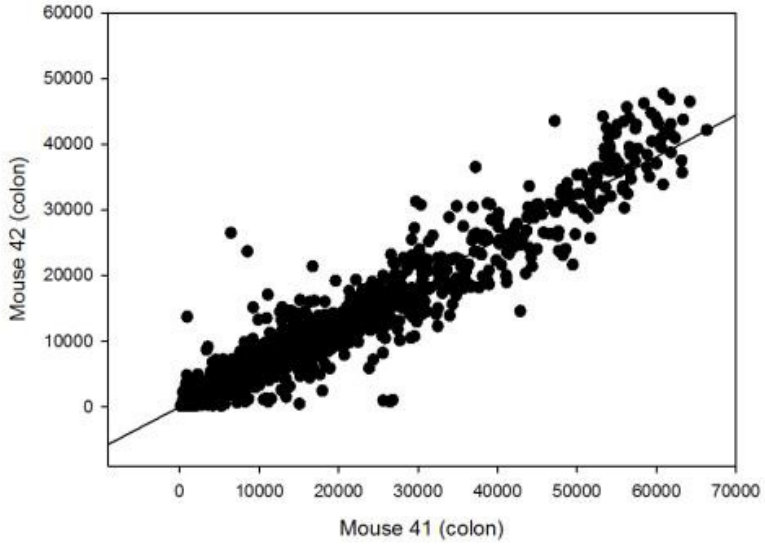
Fig.2

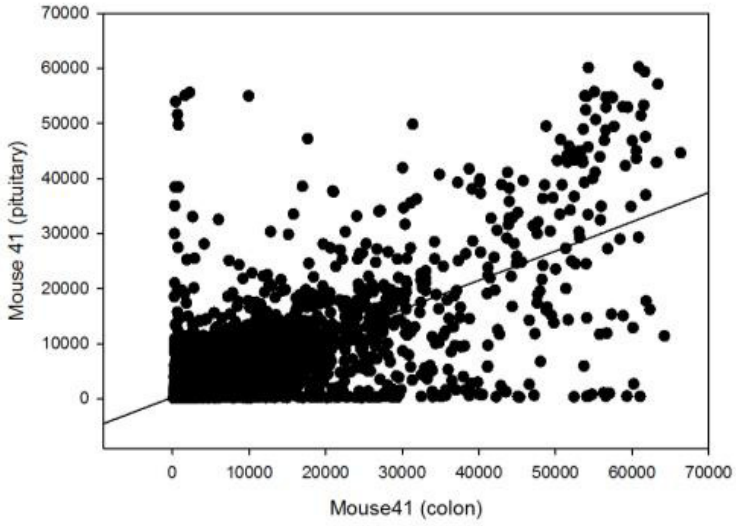


The highly degenerated reference frame, when analyzed in terms of form (correspondent to the observed correlations between landmarks) allowed to reconstruct with very high fidelity the locations of European cities as seen from Roma.

The 'degeneration' of the frame of reference was in this case a conscious choice adopted to show the power of a relatively straightforward statistical method (PCA). In biology the same 'degeneration' is a fact of nature in all the cases when we are interested in the (relatively tiny but of crucial interest) small differences induced by disease in highly invariant correlation structures. A paradigmatic case is the analysis of genome-wide gene expression patterns as measured by the microarray technology (Roden et al., 2006). Figure 3 reports the output of a classical microarray experiment. Axes correspond to two independent biological samples: in the top panel the two samples refer to the same kind of tissue (colon) for two independent rats, in the bottom panel the two samples correspond to two kind of tissues (colon and pituitary gland) of the same animal. The vector points in both panels correspond to the expression levels of around 25000 genes.

Figure 3





Looking at top panel we observe the presence of a striking invariance of shape between the two samples: the gene expression profiles are correlated near to unity (r=0.988) along four order of magnitudes (the gene expression levels reported on the axes). This order encompasses around 25000 genes (landmarks) and is a clear evidence that each tissue

corresponds to a specific balance of gene products abundance, i.e. each tissue corresponds to a specific 'shape' in the highly multidimensional gene expression space (Huang et al., 2005).
This shape invariance is totally lost in the bottom panel (r=0.235): colon (X-axis) and pituitary gland (Y-axis) correspond to two different shapes and consequently to two largely independent gene expression profiles. On the other hand, the fact to pertain to the same individual is not a sufficiently relevant order parameter to induce any correlation in the two tissues. This implies that at the level of genome-wide expression tissue communality ,there is a recognizable shape (physicists refer to this behavior as an 'attractor state' in the genome dynamics field (Huang et al., 2005) largely dominating the gene expression pattern correspondent to the tissue kind. The presence of such an overwhelming order parameter is a problem when we are interested in among individual differences (e.g. healthy vs. patients) for the same tissue: these differences are largely masked by the order imposed by tissue communality. This is exactly the same kind of problem we faced in the Europe map reconstruction starting from Latium where the role of tissue specific gene expression profile is played by PC1 (distance of European towns from the center of Latium).
Table 3 (Censi et al. 2010) reports the component loading profiles relative to a PCA as applied to a set of heart tissue samples coming from healthy and atrial fibrillation patients that must be compared with Table 2.

Table 3.

Subject codes / components	PC1	PC2
H1	0.998	-0.03
H2	0.999	-0.09
H3	0.997	-0.02
H4	0.999	-0.08
H5	0.989	-0.06
H6	0.991	-0.04
P1	0.993	0.01
P2	0.995	0.07
P3	0.996	0.05
P4	0.999	0.04
P5	0.998	0.03
P6	0.997	0.02
% expl.var.	98.4	0.3

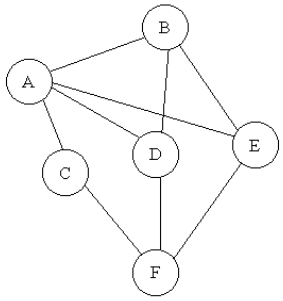
The H codes refer to healthy subjects, the P points to atrial fibrillation patients. Just like the previous case, we observe a by far major first principal component explaining 98.4% of total variability, this is a size component pointing to changes in expression that maintain the shape invariant (the tissue remain the same, even in the case of patients we are still in presence of heart tissue). Just like the previous case, the second component (PC2) is a shape component catching a relevant information about the changes in gene expression due to the disease. We can immediately appreciate the relevance of the information carried by PC2 looking at the sign of the component loadings that are negative for healthy and positive for patients. That is to say that moving along the second component does not correspond to pure size (general level of expression) changes but to shape changes: going along PC2 implies to approach the disease (increasing values of component) or healthy (decreasing values of component) status. The genes occupying the extreme poles along PC2 will be the ones endowed with the maximal explanatory power as for the pathology at hand (Censi et al., 2010, Roden et al., 2006).

2.3 Multiscale description of shape

In the previous section we demonstrated how the concept of shape as a set of correlations between variables (landmarks) allows for a straightforward analysis of multidimensional systems. The correlation matrix is the image in light of any (regular or irregular) shape: if the correlation values are all equal to unity the shape is rigidly invariant, if the correlation values can vary the shape can be modified and the entity of modifications quantified in terms of shape components.
A correlation matrix is formally identical to an adjacency matrix describing a graph in which the nodes are the landmarks and the non-null values to edges connecting the nodes. In the case of only unitary values we have an exact form, while variable values correspond to different strength of the links. Figure 4 summarizes this equivalence.

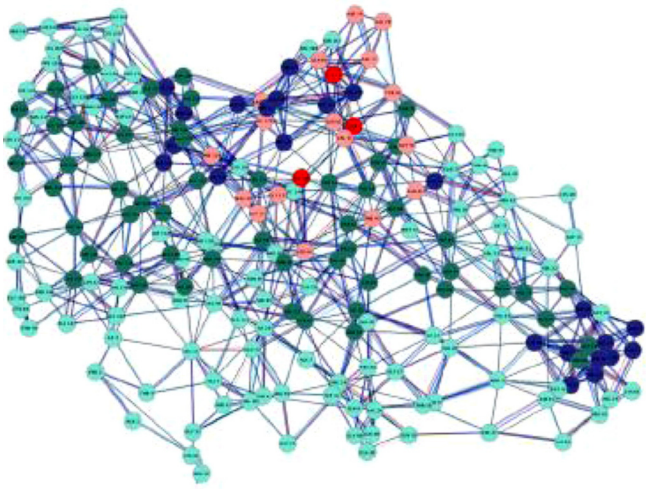
Figure 4

	A	B	C	D	E	F
A	-	1	1	1	1	
B	1	-		1	1	
C	1		-			1
D	1	1		-		1
E	1	1			-	1
F			1	1	1	-



In Fig.4, the matrix (left) is isomorphic to the graph (right) formulation and corresponds to a symmetric correlation matrix in which the trivial correlations (correlations of a variable with itself) are omitted. Graph theory is a branch of mathematics allowing for a very refined topological analysis of shapes, the equivalence between correlation and adiacency matrixes implies the possibility of a direct translation of multidimensional statistics in terms of graph theory and viceversa.
We cannot go in depth into this translation, here it is sufficient to remark that eigenvector decomposition of adiacency matrix (analogous to PCA) is the method of election for finding modules in graphs (Tasdighian et al., 2014). The point is that the adoption of graph formulation of shape (nodes=landmarks, relations=edges) allows for a very rich repertoire of graph descriptors to be applied to the studied phenomena (Dehmer and Emmert-Straub, 2014). These descriptors can be very useful to study complex biological systems naturally integrating different scales of observations and thus solving the opposition between top-down and bottom-up approaches (Juncker and Schreiber, 2008).
Figure 5 reports the contact network (graph) of a protein molecule (Di Paola et al.,2013). The nodes are the aminoacid residues (labeled with different colors according to their chemico-physical properties). Edges correspond to physical contacts in the 3D structure correspondent to a distance between 4 and 8 angstrom (distances lower than 4 Angstrom are avoided so to eliminate trivial contacts between aminoacids adjacent in the sequence, while 8 Angstrom corresponds to the upper limit of two van der Walls radiuses).

Fig.5



The above formalization allows for a direct explanation (and prediction) of relevant properties like allosteric effect (Caruso et al., 2013, Di Paola et al. 2013), cavitation (Di Paola et al., 2013), electronic transfer (Tasdighian et al., 2014).
These properties directly derive from general graph descriptors. These descriptors are located at different scales of magnification: thus we have single nodes (residues) descriptors, mesoscopic (domains) descriptors, general (entire molecule) descriptors. These descriptors follow from simple counting operations on graph. The degree of a node is a local descriptor corresponding to the number of edges starting from the node. Modularity is a mesoscopic descriptor corresponding to the proportion of within modules and total number of edges. The average degree is a general descriptor of the degree of connection of the entire graph. It is worth noting these different levels of explanation are naturally made each other dependent by the fact they all derive from the same mathematical object: thus a node with a very high degree (local level) will exert an effect on average connectivity (general level). On the other hand an elevated modularity (mesoscopic level) will influence the *betweeness* (a local descriptor, the number of shortest paths passing by a node) of those nodes connecting different modules.
This particular view on proteins shapes allows for identifying crucial elements in the structure (reductionist view) while in the same time keeping intact the global picture (systemic view). This is in my opinion the most important epistemological contribute that 'shape-focused' approach offers to scientific knowledge: to maintain alive the complexity of interactions of the studied system while in the same time allowing for a detailed characterization of its elements.

3 CONCLUSIONS

The analysis of two of the most popular methods, namely principal component analysis and graph theory, used for describing and quantifying shapes shows how the quantitative modeling of complex systems does not require sophisticated mathematics. On the contrary, the immediate intuitive flavor of the concept of shape allows for a drastic simplification of studied systems without the introduction of unjustified *a priori* hypotheses. The widespread diffusion of forms (biology and medicine are based on morphological studies since millennia) in biosphere suggests the shape-focused models as the most natural way to a sound quantification of life sciences. At this point, it is quite natural to object that the usual appreciation of forms in biology and medicine happens by means of a purely visual and picture-prone approach very far from the quantitative approaches sketched above. Notwithstanding that, this discrepancy is by a large extent purely apparent and largely dependent from the lack of basic algebraic and geometrical alphabetization of many life-science scholars. Widespread medical imaging devices, routinely used by physicians, allow for the computation of geometrical features (e.g. perimeter, area, length, width) of specific objects of interest selected by the operator. These features are computed in terms of distances between landmarks (i.e. specific points on the screen) selected by the physician. Similar procedures are routinely applied in cell biology and classical bioinformatics methods (e.g. comparison of different DNA or protein sequences) are based on the computation of distances. On the other hand, drawing box-and-arrow diagrams is the routinely adopted style of explanation in biology: these diagrams are nothing else than networks that can be naturally approached by means of graph theory.

Morphologists living in the first half of XXth century (see for example Weiss and Garber, 1952) were usually aware of the geometrical concepts and corresponding formalizations at the basis of their observations. In recent time this culture largely disappeared and it is kept alive only in relatively marginal communities as phyto-sociologists (see for example the very interesting web-page of distance-based ordination techniques cured by Oklahoma University: http://ordination.okstate.edu/). The revamping of such a quantitative consciousness allowing to generalize the naturally morphological attitude of life sciences, is now made possible by the global diffusion of computers. The promotion of such insight by sketching the bases and the far-reaching consequences of the geometrical thought was the main scope of the present paper.

4 REFERENCES

Heckman C. (1990) Geometrical Constraints on the Shape of Cultured Cells. *Cytometry,*11: 771-783.

Govindarajan, S., Recabarren, R., and Goldstein, R. A. (1999). Estimating the total number of protein folds. *Proteins: Structure, Function, and Bioinformatics*, 35(4): 408-414.

Censi F., Bartolini P., Giuliani A., Calcagnini G. (2010) A systems biology strategy on differential gene expression data discloses some biological features of atrial fibrillation. *PLoS ONE* 5 (10): e13668.

Hervè P., Derelle R., Lopez P., Plck K. et al. (2009) Phylogenomics Revives Traditional Views on Deep Animal Relationships. *Current Biology* (19): 706-712.

Berger, R. D., Akselrod, S., Gordon, D., and Cohen, R. J. (1986). An efficient algorithm for spectral analysis of heart rate variability. *Biomedical Engineering, IEEE Transactions on*, (9), 900-904.

Kauffman, S. A., and Weinberger, E. D. (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical biology*, *141*(2), 211-245.

Di Paola L., De Ruvo M., Paci P., Santoni D., Giuliani A. (2013) Protein Contact Networks: an Emerging Paradigm in Chemistry. *Chemical Reviews* 113: 1598-1613.

Di Paola L., Paci P., Santoni D., De Ruvo M., Giuliani A. (2012) Proteins as sponges: a statistical journey along protein structure organization principles. *J.Chem.Inf.Model.* 52 (2): 474-482.

Jolicoeur, P., and Mosimann, J. E. (1960). Size and shape variation in the painted turtle. A principal component analysis. *Growth*, *24*(4), 339-354.

Giuliani A., A. Colosimo, R.Benigni and J.P.Zbilut (1998) On the Constructive Role of Noise in Spatial Systems. *Physics Letters A* (247) (1,2): 47-52.

Roden, J. C., King, B. W., Trout, D., Mortazavi, A., Wold, B. J., and Hart, C. E. (2006). Mining gene expression data by interpreting principal components. *BMC Bioinformatics*, *7*(1), 194.

Huang, S., Eichler, G., Bar-Yam, Y., and Ingber, D. E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical Review Letters*, *94*(12), 128701.

Tasdighian S., Di Paola L., De Ruvo M., Paci P., Santoni D., Palumbo P., Mei G., Di Venere A., Giuliani A. (2014) Modules Identification in Protein Structures: The Topological and Geometrical Solution. *J. Chem. Inf. and Model.* (54): 159-168.

Dehmer M, Emmert-Straub F. (2014) Quantitative Graph Theory: Mathematical Foundations and Applications. CRC Press, Taylor and Francis group, New York.

Junker, B. H., and Schreiber, F. (2008). *Analysis of biological networks* (Vol. 2). John Wiley & Sons.

Caruso BL, Giuliani, A., and Colosimo, A. (2013). Allosteric transitions of proteins studied by topological networks: a preliminary investigation on human hemoglobin. *Biophysics and Bioengineering Letters*, *5*(1).

Weiss P., Garber B. (1952) Shape and Movement of Mesenchyme Cells as Functions of the Physical Structure of the Medium: Contributions to a Quantitative Morphology *Proc. Natl. Acad. Sci. USA* 38 (3): 264-280.

Recapitulation of Recurrence Theory and Practice

Charles L. Webber, Jr., Ph.D
Professor Emeritus
Department of Cell and Molecular Physiology
Loyola University Chicago, Health Sciences Division
2160 South First Avenue
Maywood, IL 60153
U.S.A.

cwebber@LUC.edu
http://homepages.LUC/~cwebber/

Keywords: Recurrence plots; recurrence quantification analysis, principle component analysis

Abstract: In broad strokes this chapter examines the roots of recurrence quantification analysis (RQA) and looks forward into its future. Much has already been written about the theory and application of recurrence methodology to real-world and complex systems. The overarching purpose of this present contribution is to introduce recurrence quantification as a unifying tool that has the timely potential of bringing together numerous and diverse fields of science under a common rubric of dynamical systems. The simple idea is that all scientists are studying "linguistic" systems that possess unique "vocabularies" and varied "grammars." But RQA can cut through these differences and encourage researchers in different research fields and scattered geographical locals to start talking seriously about the similarities of otherwise apparently diverse systems.

1. Context

The concept of recurrence has a long mathematical history that can be traced back at least to Poincare (1880) of the late 19th century and more recently to Feller (1950) of the mid 20th century. Poincare may be considered the "father" of nonlinear dynamics, whereas Feller was a mathematician focusing on probability theory. As we will see, recurrence analysis is actually a statistical tool that can "diagnose" any moving system, but we are getting ahead of ourselves. Systems can be considered as recurrent if dynamic trajectories describing those systems return exactly to (periodic) or nearby (chaotic) their starting states. According this definition, recurrent systems possess periodicities or quasi-periodicities which keep such systems intact and functioning. Any active system with strong recurrences can be considered to be in some type of steady state or quasi-steady state. However, any system exhibiting weak recurrences is either transitioning to another steady state or is moving toward dissolution (cessation). And theoretically, systems possessing no or at best unstructured recurrences are no systems at all in the classical sense, and are better understood as random processes (which nonetheless may have high-dimensional structuring that are undetermined as yet).

In the second half of the 20th century three key mathematical breakthroughs (greatly aided by the rise of digital computers) have preceded the inception of modern recurrence quantification analysis (RQA). First, in 1963 Lorenz (1963) (13,559 citations) while modeling weather systems found that small changes in initial conditions in his three-variable system led to unpredictable dynamic outcomes. This modern finding was reminiscent of Newton's three-body problem which Poincare proved had no analytical solution. Thus was born the famous Lorenz strange attractor which digital and analog computers could easily display. Besides displaying sensitivity to initial conditions, strange attractors have return trajectories that are very close but non-exact. Second, in 1981 Takens (1981) (8,206 citations) published results from his work on strange attractors in fluid turbulence. His huge contribution was the mathematical proof that dynamical systems in N dimensions could be reproduced typologically

by embedding just one of the system's variables by the method of time delays. Another way of stating this is that a single variable in a multiple-variable system can serve as the surrogate of all other coupled variables present. Third, in 1987 Eckmann, Kamphorst, and Ruelle (1987) (1,506 citations) famously merged recurrence theory with embedding theory and showed "hidden" patterns in the Lorenz system of equations. They called their new graphical tool the recurrence plot (RP) and showed how this visualization tool could be used to assess time constancy (stationarity) of dynamical systems. To them the RP was a "new diagnostic tool" as indeed it was and has proven to be multiple times over. The critical idea was that the distance between all N-point vectors derived from the input stream could be mapped into two dimensions (at i, j coordinates) using a threshold that separated small distances (close, recurrent trajectories) from large distances (far, non-recurrent trajectories). For further details of the Ruelle-Takens scenario see Eckmann (1981). So this was the state of affairs prior to the last decade of the 20th century.

Evidences of recurrence are prevalent in nature if not the entire natural world at all scales. Surely the ancients including those living in pre-history times, clearly noticed the recurrent motion of the sun which rose in the east and set in the west. Peterson (2002) paraphrases the words of King Solomon of old (Ecclesiastes 1:4):

The sun comes up and the sun goes down,
then it does it again, and again – the same old round.

If we were to put words into Solomon's mouth we could accurately say, "The sun recurs every day, but not exactly so." Indeed, annual observations of the sun shows that the points of its rising and setting shift horizontally (north and south) along the horizon. This is due, of course, to earth's axis of rotation which is tilted 23.5 degrees from the perpendicular to the ecliptic plane of revolution around the sun. As archeoastronomy shows, the ancients used these (subtle) motions of the sun to determine the proper timing for crop plantings. To us moderns, the dynamic trajectories of the sun during daylight hours are not exact (non-overlapping), but approximate each other (approximately recurrent). Indeed, to capture the full uninterrupted trajectory for an entire day, one needs to be located above the arctic in the summer as in the northern latitudes of Norway, the land of the midnight sun. Here, at midnight the summer sun "kisses" the horizon, but never sinks below it (Figure 1).

Figure 1. The midnight sun at Nordkapp, Norway. Public source:
http://en.wikipedia.org/wiki/Midnight_sun#mediaviewer/File:Midnight_sun.jpg

Dynamical systems of all flavors share one thing is common: movement in time or variations in space. For example, the following definition is as catchy as it is instructive: "If it wiggles, it's physiology; if it stops wiggling, it's anatomy!" (Webber, 2005). Physiology, of course, is a life science which is defined by a multitude of rhythms. But cessation of motion bespeaks death, leaving only the lifeless corpse. In this case, anatomy is a death science with its "dynamics" relegated solely to the spatial domain. And as an aside, the cadaver state may be the best example of a steady-state in the temporal domain due to its lack of all body motions.

The recurrence plot of a flat-line system (no temporal motion or no spatial structure) is simply one large black square. It may look like a black box, but the truth of the matter is that no hidden information is being held captive within. Rather, any flat-line system stigmatized by black box recurrences is fully known by its total lack of information and no activity. On the other hand, the presence of non-flat-line structures in any system no matter how

small will be accompanied by patchy, non-uniform recurrence plots possessing much more information. In this sense, recurrence plots can be viewed as linguistic systems converting dynamical motions and structures into visual patterns with texture (or language) that informs about the system at hand. What this means scientifically is that motions/structures in nature at the mesocopic scale as well as those at the macroscopic (large) scale to microscopic (small) and quantum (ultra-small) scales are broadcasting information than can be viewed as languages without sound as it were. Even the heterogeneous display of the cosmic background radiation in whole sky maps "screams" something instead of nothing (Fig. 2). What such diverse systems might be communicating is the subject of recurrence quantifications.

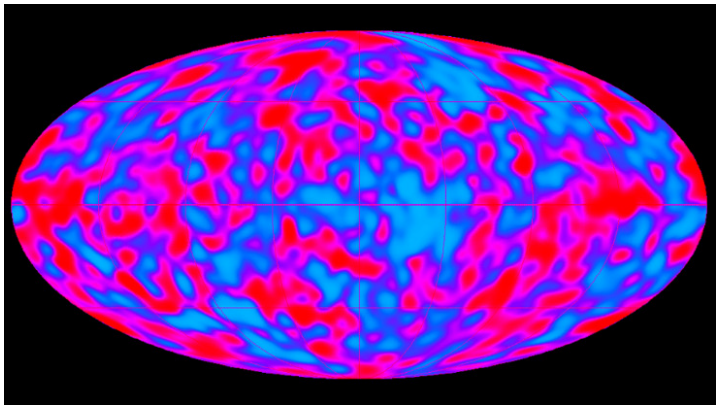


Figure 2. Whole sky map of fluctuations in the Cosmic Microwave Background (CMB) Radiation, the oldest light in universe, as recorded by the COBE spacecraft. Public source: http://en.wikipedia.org/wiki/Cosmic_Background_Explorer#mediaviewer/File:COBE_cmb_fluctuations.gif

2. Codings

Matrices are ideal structures in which meaningful information can be stored as well as retrieved. One excellent example is the Quick Response Code which is nothing more than a two-dimensional barcode. QRs, as they are called, appear as light and dark squares distributed within a large square box (or window) with three of the four corners designated with identical target patterns (for alignment). As shown in Figure 3, QR patterns look exactly like cross-recurrence plots (non-symmetrical) which contain meaningful information (text, WebSite links, etc.). The patterns can be conveniently captured via any cell phone camera and decoded with any QR apt. The big idea being conveyed here is that dynamical patterns in nature possess recurrence structures that can be projected or revealed as recurrence plots (like QR codes) that await reading (like QR apt). As we will see, recurrence plots provide the patterns, but recurrence quantifications provide the interpretations.

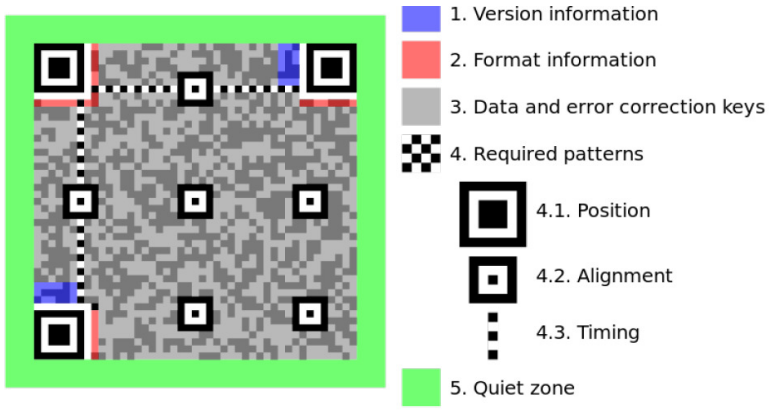


Figure 3. Quick Response coding used to embed detailed information and Web Page links.
Public source: http://en.wikipedia.org/wiki/QR_code#mediaviewer/File:QR_Code_Structure_Example_3.svg

3. Calculations

Digital computers and matrix mathematics have made it possible to generate recurrence plots of dynamical systems in time or space with ease. Signals in the temporal or spatial domain can first be embedded by the method of time delays (Takens, 1981) and then mapped to recurrence space as recurrence plots (Eckmann et al., 1987) to yield beautiful two-dimensional plots (curious QRs at best). But just what are the meanings of these displays? Zbilut and Webber (1992) (514 citations) and Webber and Zbilut (1994) (769 citations) tackled this problem by introducing recurrence quantification analysis (RQA) into both the physics literature and physiological literature, respectively. Simply put, Webber and Zbilut defined five recurrence variables that were unambiguously extracted (excluding personal bias) from the recurrence plot including: recurrence (REC, density of recurrent points); determinism (DET, diagonal line structures); diagonal maximum (DMAX, longest diagonal line); entropy (ENT, distribution of diagonal line structures); trend (TND, homogeneity of recurrent points). To these five variables, in 2002 Marwan et al. (2002) (409 citations) added three more RQA variables: laminarity (LAM, vertical line structures); vertical maximum (VMAX, longest vertical line); trapping time (TT, average vertical line length).

A major 14-chapter volume edited by Webber and Marwan (2014) has just been published that describes all the mathematical details and practical implementations of recurrence plots and recurrence quantifications. And a second volume follows closely behind (Marwan, Riley, Giuliani and Webber, 2014). So there is no need to reiterate the facts in the restricted space here. Suffice it to say that each extracted recurrence variable has a dynamical interpretation that imparts meaning not only to the recurrence plots, but to the input signals themselves. Metaphorically, RQA can be viewed like a Swiss army knife with eight different purposes (Fig. 4). Thus, recurrence indicates the density of recurrence structures present, determinism reveals the level of rule-obeying structures in operation, maximal diagonal line reports on the chaoticity of the signal, entropy captures the complexity of the signal, trend is sensitive to the stationarity of the system, maximum vertical line is sensitive to the stochasticity of the signal, laminary reveals the level of rule-defying dynamics, and trapping time is an average stochasticity measure. The experimenter coming upon a difficult and complex dynamic now has a sophisticated 8-pronged tool with which to approach the problem. Maybe one or two prongs are sufficient to solve the problem; maybe all eight prongs are required to unlock the mystery (if we can call it that).

Figure 4. Swiss army knife metaphor for the 8 recurrence variables, 5 from Webber and Zbilut (red labels) and 3 from Marwan (blue labels).

The original RQA software was programmed by Webber which is now in its 14th version (Webber, 2012). The programming language used has always been the C language and complied programs still run under the emulated and antiquated DOS system (e.g. DosBox). Officially DOS stands for Disk Operating System, but Webber (2009) dubbed it as "Dinosaur or Something" years ago. Since URL hits were first counted starting on August 22, 1996 almost 30,000 hits have been registered to Webber's site. But there are some interesting, hidden facts about the development of this software. For example, initial program forays into the world of recurrence were named

Recurrence Plot Analysis (RPA). This was a steep learning curve of trial and error, find-tuning and error correction. Early programs were distributed by snail-mail to interested users on 3.5 inch diskettes (freeware).

Despite all cautions, in the mid 1990's I discovered a heart-breaking error in the software which I now call a confession. Close examination of recurrence plots displayed antiparallel lines even what the embedding dimension was greater than 1. Such palindromic structures were simply impossible for embedded time series in linear time. The problem was quickly isolated to an incorrect calculation of vector distances. That is, what was believed to be the distance between vector pairs (VD) was actually the difference in vector magnitudes (VM)! Here, VM and VD are equivalent only for embedding dimensions (ED) of 1. The misapplied and incorrect formula for VM (Eq. 1) was immediately replaced by the appropriate formula for VD (Eq. 2), and warning notes were sent out to all known RPA users. Subsequently, the entire software package was renamed as RQA. In the meantime numerous recomputations on previously published time series were done to assess the damage done. Fortuitously, the results confirmed that the conclusions of those studies were basically correct. Thus in one key example, muscle fatigue was detected earlier by both RQA vector distances (Fig. 2.9 in Webber and Zbilut, 2005) and RPA vector magnitude differences (Fig. 5 in Webber et al., 1995) both compared against the linear Fast Fourier Transforms (FFT).

$$VM = ABS[SQRT(P_{i1}^2 + P_{i2}^2 + \ldots + P_{iED}^2) - SQRT(P_{j1}^2 + P_{j2}^2 + \ldots + P_{jED}^2)] \qquad (1)$$

$$VD = SQRT[(P_{i1}-P_{j1})^2 + (P_{i2}-P_{j2})^2 + (P_{i3}-P_{j3})^2 + \ldots + (P_{iN}-P_{jED})^2] \qquad (2)$$

The next recurrence software to be freely distributed was Visual Recurrence Analysis (VRA) as programmed in C++ for Windows computers by Kononov in 1996. His latest version of VRA was 4.9 published a decade later (Kononov, 2006). This software is very easy to use and focuses on unthresholded recurrence plots. Still, VRA can compute recurrence variables with exact correspondence to RQA ensured by the fact that Webber shared RQA code with Kononov. Kononov is one impressive programmer who works in financial markets.

The third software package offered for free use was Cross Recurrence Plots (CRP) as programmed in M-files for MatLab platforms by Marwan (2013). CRP is now up to version 5.17 in 2013. Marwan has enjoyed up to almost 24,000 hits to his CRPtoolbox since 2005. MatLab, of course, is a very popular platform among different scientific disciplines. Marwan also shared his M files to users can understand how computations are made. He is another clever programmer.

As a last thought on calculations, one of my longstanding pet peeves as a professional reviewer for scientific journals has been the misuse of the term parameter. Parameters, of course, are adjustable constants. Consider the famous logistic difference equation of May (1976) which possesses three specific components: the constant "1"; the changing variable "x_n"; the parameter or adjustable constant "a."

$$x_{n+1} = a \bullet x_n \bullet (1 - x_n) \qquad (3)$$

Here "a" is a tuning parameter which when set to low values ($a < 3.0$) results in monotonic dynamics in iterated variable "x_n". With higher fixed values of "a" ($3.1 < a < 3.5$) the dynamics of "x_n" become periodic with successive period doublings. And with the highest values of "a" ($36. < a <= 4.0$) the dynamics of "x_n" are in the chaotic mode (with some narrow periodic windows of calm). Recurrence properties of the logistic equation were first described by Trulla et al. (1996), unfortunately with the incorrect calculation of distances between vector pairs.

Looking back on my early publications I note the misnaming of physiological variables as parameters in in the titles of two abstracts (Pleschka et al., 1975; Speck and Webber, 1978) and the text of at least two full publications (Webber and Peiss, 1975, 1979). But as "misery loves company," Eckmann et al. (1987) made the same blunder when they wrote in their famous recurrence plot paper: "In recent years a number of methods have been devised to compute dynamical parameters from time series." Again, what these mathematicians refer to as "parameters" are really "variables." Just to be clear on this point, to date there are 8 recurrence variables in use and they should never be construed as dynamical parameters. Fixed parameters are adjustable but are held constant during any calculations. But the question is how does one set these parameters properly? This challenging issue is now discussed.

4. Challenges

There are two levels of difficulty surrounding the proper implementation of recurrence plots and recurrence quantifications which have been addressed thoroughly elsewhere (Webber and Marwan, 2012; Marwan, Riley, Giuliani and Webber, 2014, Webber and Zbilut, 2005; Marwan et al. (2007). First is the challenge of setting the RQA parameters correctly; and second this is the challenge of interpreting the recurrence plots and their quantifications wisely. On the input side of things, there are a full 8 recurrence parameters each of which must be set properly and this is no small task. These 8 adjustable parameters include: embedding dimension (EMB), delay (DLY), window (WIN), norm (NORM), rescaling (RESCALE), radius (RAD), line (LINE) and shift (SHIFT). In the RQA software of Webber (2012) these parameters are entered sequentially, one by one, and illegal combinations are prohibited. The two most important parameters are the embedding dimension and radius (threshold). The embedding is important because it represents the dimensionality into which the dynamic system under study must be contained. In addition, real-world systems are noisy and noise inflated the apparent dimension (Parker and Chua, 1989). To estimate the best embedding dimension (it is better to over-estimate than under-estimate the dimension) one can apply the false nearest neighbor approach (Kennel et al., 1992).

The second parameter that can be tricky to set is the radius which determines the boundary between vector distances ($D_{i,j}$) that are defined as being recurrent ($D_{i,j}$ <= radius) or non-recurrent ($D_{i,j}$ > radius). An unthresholded recurrence plot employs no radius (or sets radius > max $D_{i,j}$), but color codes each recurrent point depending upon the distance at that point as was classically implemented in VRA software (Kononov, 2006). Mathematically speaking, insertion of the radius discriminator actually converts the distance matrix (table of distances) into the recurrence matrix (table of recurrent 1's and non-recurrent 0's). Black and white recurrence plots are simply visualizations of the recurrence matrix at correspondent locations (black points for the 1's; white points for the 0's). If the radius is set too low, the recurrence density can be too sparse for statistical utility. If the radius is set too high, paired vectors that are "far" from one another are nonetheless declared "close" but incorrectly so. This is not the place to give all the ins and outs of how to practically solve these issues. But one piece of advice may prove useful. Pretend that you are a photographer and that the embedding dimension is the shutter speed and the radius is the F-stop. To "take a photo" of your dynamic why not "bracket your exposures"? That is, one can study the same dynamic on multiple passes using different combinations of embedding dimensions (e.g. 5 10, 15, 20) and radii (1%, 2%, 5%, 10%, etc.). Then the results can be studied to see which output combinations give results that are most robust. This is moving from theory to practicality, a necessary and sufficient practice of scientists no matter what their field of inquiry.

Now on the output side of things, there is a set of 8 recurrence variables that are computed for each recurrence window. As inferred above, these output values critically depend upon the parameter settings. This fact alone drives the wisdom of trying different combinations of parameters (within reason) on the same input data set to observe what works best. Back in 1994 Webber and Zbilut (1994) demonstrated how the recurrence variables could be rendered as dynamic variables by introducing the system of moving windows. Here is where the shift parameter comes into play. After computations were completed within one window, the frame of reference could be shifted into the future to capture a second set of computations. Large shifts or steps could be computed quickly, but at the loss of resolution. Small shifts (e.g. with window steps as low as 1) the highest resolution could be acquired, but at expense of long computation times.

Staying with the linguist theme, let us now consider the famous children's poem by Dr. Seuss, "Green Eggs and Ham" (Geisel, 1960). This example, by the way, has proven to be remarkable for introducing newcomers to the world of recurrence plots and quantifications as illustrated by Webber and Zbilut (1996). This rhythmical poem consists of 812 words, but the vocabulary is restricted to only 50 different words. This means that the same few words must be reused to build the entire poem. In recurrence terminology, reused words are simply recurrent words. Three steps are taken to compute recurrence quantifiers for this poem: first, encode each new word in the poem with a unique integer (1-50) to construct a vector of integers (N = 812); second, compute recurrence quantifications (program RQE) selecting an embedding dimension of 3 (short 3-word vectors), a radius of 0 (only identical 3-word sequences recur), a window of 80 (one tenth length of poem) and a shift of 1 (maximum resolution). Third, plot the 5 traditional recurrence variables over the 731 sliding-window epics as is shown in Figure 5 (the 3 variables of LAM, VMAX and TT show no dynamics in this example). Close examination of variables REC, DET, DMAX, ENT and TND reveals five unique rhythmical structures. The structuring of the poem can be completely destroyed by random shuffling of the words order. All words are represented, but their unique sequencing (along with the

story they tell) is lost. In this case, the recurrence variables are greatly changed (e.g. decreases in REC and DET, etc.) suggesting that shuffled points in any input series may serve as a control for the native ordering of those same points.

It is important to remember that these five variables are all being computed from the same input series, but each from its own perspective. From this illustration it is good to think of the 5 (nay 8) recurrence variables as separate "observers" of the same "event" from their own particular point of view. This is exactly what happens when 8 witness attempt to describe the details of the same car accident. Each has a truth to tell, but only within individualized frameworks, contexts and biases.

Analyzing text by matching word for word (embedding dimension = 1) or chain of words for chain of word (embedding dimension > 1) is based on similarity of terms. Recently, Argus et al. (2012) have introduced what they refer to as conceptual recurrence plots. Here recurrence similarity is not based on exact words (or letters) but on conceptual similarities within speech patterns. This approach is particularly useful when analyzing a complex conversation between two subjects, like doctor and patient or interviewer and interviewee.

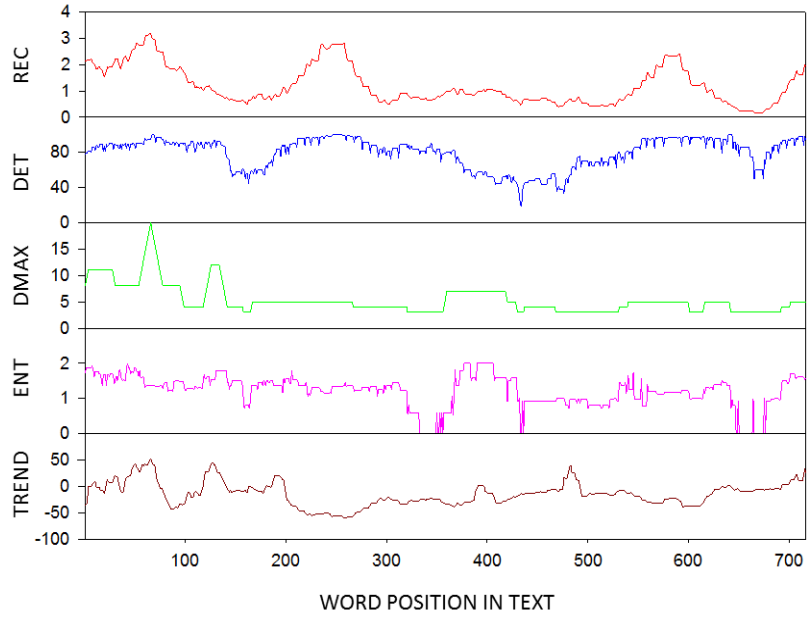


Figure 5. Recurrence quantification variables computed within 731 moving windows, each 80 words long, sliding over the entire poem, "Green Eggs and Ham." This poem has 8212 words and recurrences are scored only on exactly matching 3-word sequences (e.g. "I am Sam" recurs with "I am Sam" but not with "Sam I am").

5. Connectome

The linguistic motif easily extends to the biological world of genomics and proteomics. Deoxyribonucleic acid (DNA) is the language of life and consists of only four base pairs (A for adenine, C for cytosine, G for guanine, T for thymine). Unique triplets (24 permutations) of bases (embedding of 3) code for each of the 20 naturally occurring amino acids in the construction of linear proteins on the ribosomes. These protein strings are subsequently folded into 3-dimensional structures (knots?) on the endoplasmic reticulum. For example, the recurrence structure of the entire genome of bacteriophage lambda (48,502 nucleotides) has been studied with a 1000-point sliding window (Webber and Zbilut, 1998). Likewise, the recurrence structure of entire proteins can similarly be examined using recurrence quantifies. However, beyond coding individual amino acids in their native sequence by arbitrary integer coding, actual hydrophobicity values for each protein can be substituted, allowing non-zero RQA radii to be used. Moving into the real physical domain has been exceptional fruitful in describing different nonlinear aspects of proteins (Zbilut et al., (2004).

The entire natural world is replete with rhythms, each of which is conveying a language with meaning. The tides rise and fall, weather patterns swirl around the globe, earthquakes occur aperiodically in different locals. In physics there is the entire electromagnetic spectrum which encompasses concurrent oscillations in numerous waveforms at different frequencies. Indeed, string theory proposes that the fundamental basis of physical reality rests on vibrating, one dimensional strings each with a length, but no width (dimensionless). But more than theory is the periodic table of the elements which cluster according to physical properties attributed to the unique mixing and matching subatomic particles of which they consist (protons, neutrons, electrons, quarks, etc.). So no matter which system one is examining, each and every system is amenable to recurrence analysis. Thus, to study "wiggling" dynamics implies that recurrence analysis is a linguistic interpreter.

Another linguistic system of great importance is the central nervous system of the human brain. If one thinks the estimated number of neurons at 86 billion is staggering (Azevedo et al., 2009), consider the 100 trillion synaptic connections with great diversity among all those neurons (O'Rourke et al., 2012)! Initial attempts to map the somatosensory homunculus have yielded static pictures of the human brain. But much more recently capturing the topological and spatial organization of the brain in terms of synaptic connections or connectomes gives a much more dynamic view of the brain. Indeed, one gets a better feel for the flow of information one brain regions to another (Kaiser, 2011). One researcher boasts that we are more than our genome; we are our connectome (Seung, 2012). Figure 6 is a functional connectivity matrix of the human brain from The Human Connectome Project sponsored by the National Institutes of Health (NIH). The matrix presents as a correlation matrix similar to an unthresholded recurrence plot. It is now posited that adding recurrence quantifications to brain activities would be a powerful contribution to understanding the functional rhythms of the brain in time and space. The gross electroencephalogram has been studied in this manner (Thomasson et al., 2002).

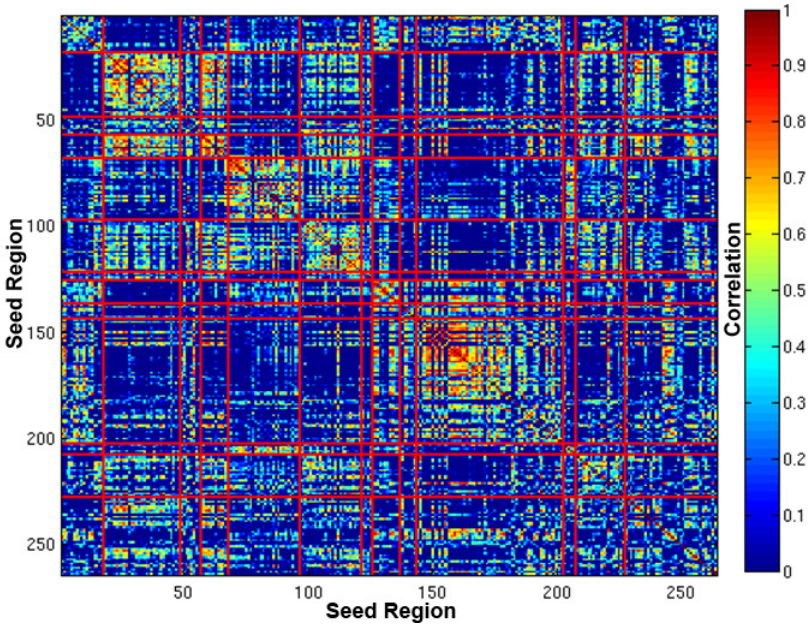


Figure 6. Functional connectivity matrix, fictive of an unthresholded recurrence plot.
http://www.neuroscienceblueprint.nih.gov/connectome/

6. Combinations

The case has been built herein that dynamical systems in time and space possess rhythms or patterns, respectively, that are amenable (ripe) for recurrence analysis. RQA is a nonlinear tool which makes no assumptions regarding the type, length or even quality of the data. Input data need not be Gaussian distributed, can be relatively short (as low as 30 scalars in the vector), can be either nonstationary or stationary, and corrupted by noise (Zbilut et

al., 1998). These things stated, RQA is not an end-all or be-all of tools. This is why the concept of mixing and matching methodologies provide the best approach to problem solving and system understanding.

To illustrate this key point, I would like to introduce what is now called the "linguistiscope" as shown in Figure 7. This is a very unique instrument that combines the power of the RP and RQA with Principle Component Analysis (PCA). It is very different from the "retrospectroscope" of Comroe (1977) which is of historical importance. The design this linguistiscope is both like a telescope (magnifying distant objects) and a microscope (magnifying small objects) using three lenses. The user points this device at his/her system of interest which most likely is multidimensional, nonlinear, and noisy. The first lens (RP) collapses the N-dimensional signal down into a two-dimensional and thresholded recurrence plot. The second lens (RQA) expands the compressed signal back up into an eight-dimensional vector of recurrence quantifiers (for each moving window). The third lens (PCA) compresses the matrix of RQA variables into the first three principle components (PC1, PC2, PC3) which typically capture 95% of the variance. This third lens frees the investigator from having to pick and choose which one RQA variable is most sensitive for the dynamic at hand. When all is said and done, the observer "eyes" the data graphically in 3-D plots (PC1 vs. PC2 vs. PC3) or 2-D plots (PC1 vs. PC2 or PC1 vs. PC3 or PC2 vs. PC3). By this means, different sets of nonlinear experimental data can be sorted and classified linearly. This idea of "marrying" RQA and PCA and proof of concept on it power in analyzing complex systems can be read about elsewhere (Zbilut et al., 1998; Giuliani et al., 2001).

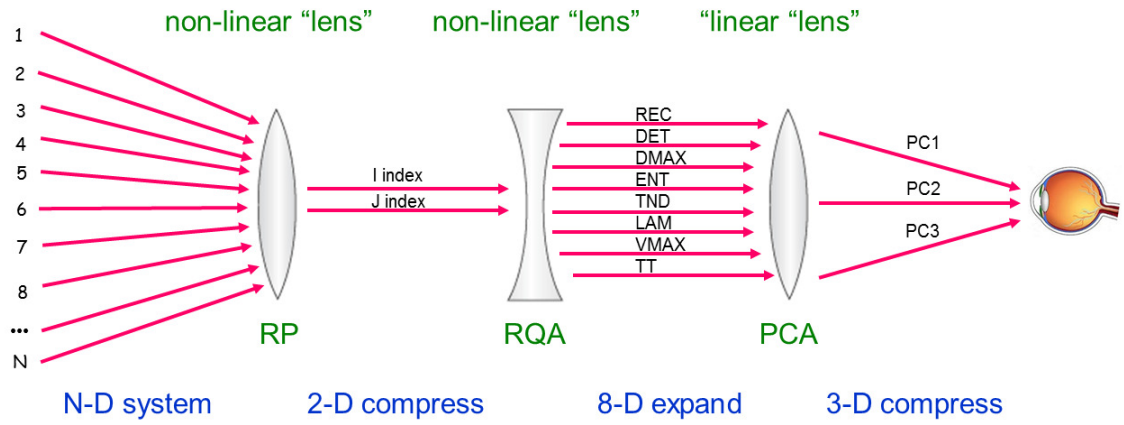


Figure 7. The Linguistiscope combines the advantages of recurrence plots (RP) and recurrence quantification analysis (RQA) with principle component analysis (PCA) to focus in on dynamical systems of any scale. Data flowing forth from these structures are treated as linguistic strings which are communicating dynamical details of the behavior of any system on any scale in time and space.

7. Conversations

It is intentional that this chapter ends not with conclusions, for the conversations about recurrence plots and recurrence quantification are continuing and open-ended. With some scrutiny, it can be observed that experimenters and theoreticians representing a vast variety of scientific disciplines and fields are basically asking rather similar questions, but do so by employing concepts and vocabularies that are very different and foreign to one another. What is suggested is that RQA is a unifying tool that can bring together scientists of very different persuasions under the unity umbrella of a common recurrence vocabulary. To the extent that dynamical inputs from any field are understood to be linguistic systems (multilingual albeit language generating), recurrence analyses can uncover similar patterns or dynamics in systems that are otherwise very different. For example, personal interactions in psychology (Shockley and Riley, 2014) are not all that different from coupling properties of strange attractors in mathematics (Chelidze and Matcharashvili , 2014). Demonstration that RQA brings researchers together is seen in

the continuing success of bi-annual International Symposia on Recurrence Plots: Potsdam (2005); Siena (2005); Montreal (2009); Hong Kong (2011); Chicago (2013); Grenoble (2015 projected).

Recurrence plots are very beautiful to look at, so much so that initiates can stall at this gestalt-feeling level and miss the deeper meanings as carried in the recurrence quantifiers. It is critically important to move from the pretty qualitative images to the necessary quantitative descriptions. This said, as one studies recurrence plots, different patterns can be observed which may stimulate the definition of new recurrence variables. As an example, take a look at the pixelated painting in Figure 8. Here the artist depicts large-scale curved structures which can only be appreciated by observing the full collection of 2,304 pixels, but not by any of the 8 RQA descriptors. Curved structures have indeed been observed in recurrence plots (Facchini and Kantz, 2007) which can be attributed to nonstationary signals (Facchini et al., 2005) or synchronization phenomena of two time series (Marwan et al., 2002). But how can such structures be defined mathematically? This challenge is left for future researches to solve.

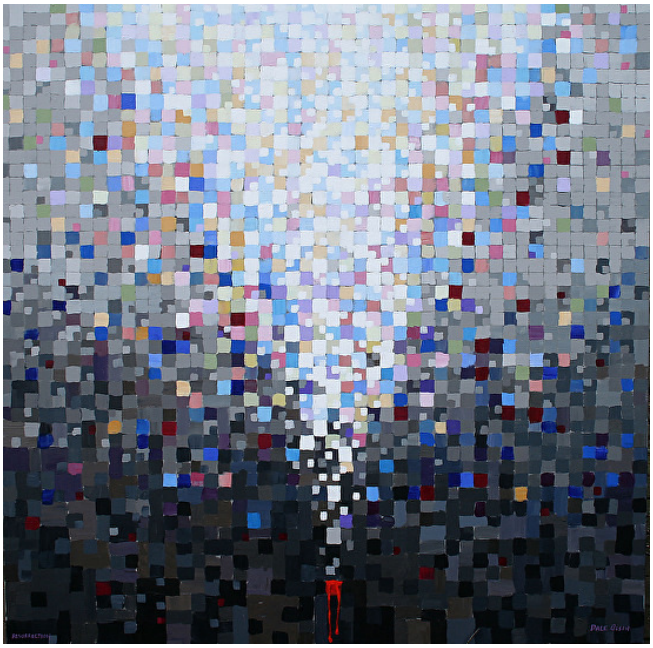


Figure 8. The painting "Resurrection" by Dale Olsen consisting of 2,304 pixels arranged in a square matrix [48, 48]. Used with written permission from the artist: http://daleolsen.fineartstudioonline.com/works

One final thought is offered to unify researchers and keep the conversations going across the greater scientific community. One experimental design that seems to ubiquitous across disciplines is the desire/need to have early warnings for catastrophic events. In medicine it would be of great advantage to be able to forecast heart attacks or seizures. In industry it would be fiscally responsible to predict system failures compromising product qualities. In meteorology it would be lifesaving to warm the masses of impending severe storms such as tornados, monsoons, floods, etc. In the old days it is said that farmers needed not look to the sky to sense an approaching thunder storm, but rather to watch the changing behavior of their barnyard chickens (which purportedly sense the electrostatic charges in the atmosphere). In more recent times animal behaviors have been found to be different and unusual prior to earthquakes (Buskirk et al., 2010). Might the "linguistiscope" be seen as a better "chicken" sensitive to that which is beyond human sensitivities?

To illustrate this important point, let us pursue the earthquake theme. In the earth sciences earthquake prediction is as critical as it is illusive. How important it is to understand the movement of the earth's crust again to warm and evacuate large populations of people and get them out of harm's way. Earthquakes are indeed being studied using recurrence strategies on seismographic signals (Chelidze and Matcharashvili, 2014). Figure 8 outlines

one way to apply RQA to earthquake dynamics, but this example extends to any early-forecast system. The idea is this. Seismographs measure the tremors in the earth's crust. It is possible that there are subtle changes in the recorded signals preceding "the big event" that can be magnified by RQA and give ample warning. The experimental design would simply be to continuously feed the seismographic data string(s) into the Linguistiscope and analyze the dynamics occurring within each moving window. Hopefully, early detection would be realized.

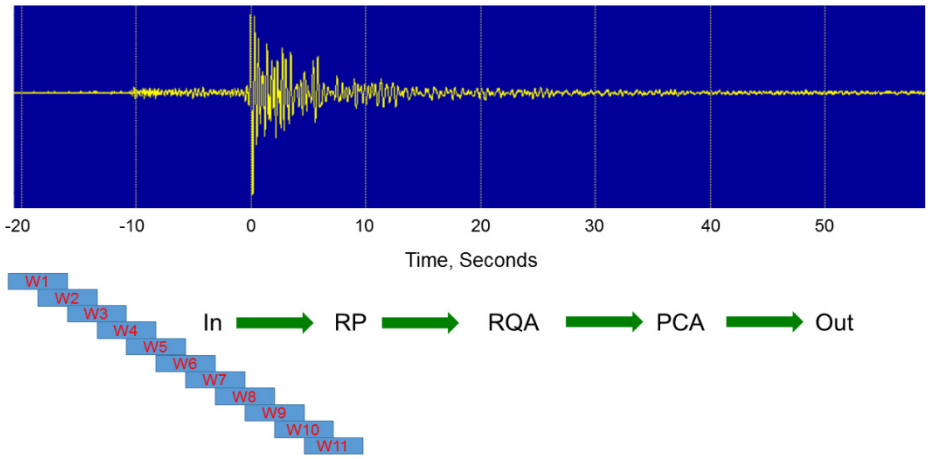


Figure 9. Schema for the detection of pre-catastrophic event by sliding window recurrences.

So we leave this discussion at this point. These is so much more that could be said and that will be said in the years and decades to come. The field of recurrence analysis may be very young, but it promises a bright future of applications as yet not realized.

8. References

Angus, D, Smith, A., and Wiles, J. (2012). Conceptual recurrence plots: Revealing patterns in human discourse. *IEEE Trans. Visualization Comp. Graphics* 18:988 - 997.

Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti. R. E, Leite, R. E., Jacob Filho, W, Lent, R, and Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* 513:532 - 41.

Buskirk, R. E., Frohlic C., and Latham, G. V. (2010). Unusual animal behavior before earthquakes: A review of possible sensory mechanisms. *Rev. Geophysics* 19:247 - 270.

Comroe, J. H., Jr. (1977). *Retro Spectro Scope: Insights into Medical Discovery.* Von Gehr Press, Menlo Park, CA.

Eckmann, J.-P. (1981). Roads to turbulence in dissipative dynamical systems. *Rev. Mod. Physics* 53:643 - 654.

Eckmann, J.-P., Kamphorst, S. O., and Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhys. Lett.* 4:973 - 977.

Facchini, A and Kantz, H. (2007). Curved structures in recurrence plots: the role of the sampling time. *Phys. Rev. E* 75: 036215.

Facchini, A., Kantz, H., and Tiezzi, E. (2005). Recurrence plot analysis of nonstationary data: The understanding of curved patterns. *Phys. Rev. E* 72, 021915.

Feller, W. (1950). *An Introduction to Probability Theory and Its Applications.* John Wiley and Sons, New York, Volume 1.

Geisel, T. S. (pseud. Dr. Seuss) (1960). *Green Eggs and Ham.* Beginner Books, Random House, Inc., New York.

Giuliani, A., Colafranceschi, M., Webber, C. L., Jr., Zbilut, J. P. (2001). A complexity score derived from principal components analysis of nonlinear order measures. *Physica A* 301:567 - 588.

Kaiser, M. (2011). A tutorial in connectome analysis: Topological and spatial features of brain networks. *NeuroImage* 57:892 - 907.

Kennel, M. B., Brown, B., and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev.A* 45, 3403-3411.

Kononov, E. (2006). Visual Recurrence Analysis (VRA) software for Windows, version 4.9 available at http://visual-recurrence-analysis.software.informer.com/4.9/

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *J. Atmos. Sci.* 20:120 - 141.

Marwan, N. (2013). Cross Recurrence Plot Toolbox (CRP) software for MatLab, version 5.17 available at http://tocsy.pik-potsdam.de/CRPtoolbox/

Marwan, N., Riley, M., Giuliani, A., and Webber, C. L., Jr. (eds.) (2014). *Translational Recurrences: From Mathematical Theory to Real-World Applications*. Understanding Complex Systems, Springer International Publishing, Switzerland.

Marwan, N., Romano, M. C., Thiel, M., and Kurths, J (2007). Recurrence plots for the analysis of complex systems. *Phys. Rep.* 438:237 - 329.

Marwan, N., Thiel, M., and Nowaczyk, N. R. (2002). Cross Recurrence Plot Based Synchronization of Time Series. *Nonlin. Processes Geophys.* 9:325 – 331.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., and Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart rate variability data. *Phys. Rev E* 66:026702.1-026702.8.

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature* 261:459 - 467.

O'Rourke, N. A., Weiler, N. C., Micheva1, K. D., and Smith, S. J. (2012). Deep molecular diversity of mammalian synapses: Why it matters and how to measure it. *Nature Rev. Neurosci.* 13:365 - 379.

Parker, T. S, and Chua, L. O. (1989). *Practical Numerical Algorithms for Chaotic Systems.* Springer-Verlag, New York.

Peterson, E. (2002). *The Message: The Bible in Contemporary Language.* NavPress, Colorado Springs. p. 1164.

Pleschka, K., Riedel, W., and Webber, C. L., Jr. (1975). Correlation of ventilatory parameters and patterns of phrenic motoneuron discharge in the resting and panting cat. *Pflügers Arch.* 359:R50.

Poincaré, H. (1890). Sur la probleme des trois corps et les équations de la dynamique. *Acta Math.*, 13:1 - 271.

Seung, S. (2012). *Connectome: How the Brain's Wiring Makes Us Who We Are.* Houghton Mifflin Harcourt Publishing Company, New York.

Shockley, K. and Michael Riley, M. A. (2014). Interpersonal couplings in human interactions. In: *Recurrence Quantification Analysis: From Theory to Best Practices.* Webber, C. L., Jr. and Marwan, N. (eds.). Springer International Publishing, Switzerland, chapter 14, pp. 61-95.

Speck, D. F. and Webber, C. L., Jr. (1978). Effect of thoracic dorsal rhizotomy on respiratory parameters in quiet breathing. *Fed. Proc.* 37:368.

Takens, F. (1981). Detecting strange attractors in turbulence. In: *Lecture Notes in Mathematics, Vol. 898, Dynamical Systems and Turbulence, Warwick 1980*. Rand, D. and Young, L.-S. (eds.), Springer-Verlag, Berlin, pp. 366-381.

Tamaz Chelidze and Trimuraz Matcharashvili (2014). Dynamical patterns in seismology. In: *Recurrence Quantification Analysis: From Theory to Best Practices.* Webber, C. L., Jr. and Marwan, N. (eds.). Springer International Publishing, Switzerland, chapter 10, pp. 285-328.

Thomasson, N., Webber, C. L., Jr., and Zbilut, J. P. (2002). Application of recurrence quantification analysis to EEG signals. *Int. J. Comp. Appl.* 9: 1-6.

Trulla, L. L., Giuliani, A., Zbilut, J. P., and Webber, C. L., Jr. (1996). Recurrence quantification analysis of the logistic equation with transients. *Phys. Lett. A* 223:255 - 260.

Webber, C. L., Jr. (2005). The meaning and measurement of physiologic variability. *Crit. Care Med.* 33:677 - 678.

Webber, C. L., Jr. (2009). DOS (Dinosaur or Something) implementation of Recurrence Quantification Analysis. In: *Third International Symposium on Recurrence Plots*. Montreal, Canada.

Webber, C. L., Jr. (2012). Recurrence Quantification Analysis (RQA) software for DOS, version 14.1 available at http://homepages.luc.edu/~cwebber/

Webber, C. L., Jr. and Marwan, N. (eds.) (2014). *Recurrence Quantification Analysis: From Theory to Best Practices.* Understanding Complex Systems, Springer International Publishing, Switzerland.

Webber, C. L., Jr. and Peiss, C. N. (1975). Interaction of respiratory cell discharge patterns and spontaneous respiratory rate. *Amer. J. Physiol.* 228:1384 - 1392.

Webber, C. L., Jr. and Peiss, C. N. (1979). Pentobarbital-induced apneusis in intact, vagotomized and pneumotaxic-lesioned cats. *Respir. Physiol.* 38:37 - 57.

Webber, C. L., Jr., Schmidt, M. A., and Walsh, J. M. (1995). Influence of isometric loading on biceps EMG dynamics as assessed by linear and nonlinear tools. J. Appl. Physiol. 78:814 - 822.

Webber, C. L., Jr. and Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol.* 76:965 - 973.

Webber, C. L., Jr., and Zbilut, J. P. (1996). Assessing deterministic structures in physiological systems using recurrence plot strategies. In: *Bioengineering Approaches to Pulmonary Physiology and Medicine*, M. C. K. Khoo, M. C. K. (ed.). Plenum Press, New York, pp. 137 – 148.

Webber, C. L., Jr. and Zbilut, J. P. (1998). Recurrent structuring of dynamical and spatial systems. In: *Complexity in the Living: A Modelistic Approach.* Colosimo, A. (ed). Proc. Int. Meet., Feb. 1997, University of Rome "La Sapienza," pp. 101-133.

Webber, C. L., Jr. and Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In: *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences.* Riley, M. A, and Van Orden, G. (eds.), chapter 2, pp. 26 - 94.

Zbilut, J. P., Giuliani, A., and Webber, C. L., Jr. (1998). Recurrence quantification analysis and principle components in the detection of short complex signals. *Phys. Lett. A.* 237:131 - 135.

Zbilut, J.P., Giuliani, A., and Webber, C.L., Jr. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Phys. Lett. A* 246: 122-128.

Zbilut, J. P., Mitchell, J. C., Giuliano, A., Colosimo, A., and Marwan, N., Colafranceschi, M., and Webber, C. L., Jr. (2004). Aggregation propensity of proteins quantified by hydrophobicity patterns and net change. In: *Complexity in the Living: A Problem-Oriented Approach.* Benigni, R., Colosimo, A., Giuliani, A., Sirabella, P. and Zbilut, J. P. (eds.). Proc. Int. Meet., Sept. 2004, University of Rome "La Sapienza," pp. 136-151.

Zbilut, J. P. and Webber, C. L., Jr. (1992). Embeddings and delays as derived from quantification of recurrence plots. *Phys. Lett. A* 171:199 - 203.

ON THE ABUNDANCE OF INTRINSICALLY DISORDERED PROTEINS IN THE HUMAN PROTEOME AND ITS RELATION TO DISEASES: THERE IS NO ENRICHMENT

Antonio Deiana
Physics Department, Sapienza University of Rome, P.le A. Moro 2, 00185 Rome, Italy
antonio.deiana@roma1.infn.it

Andrea Giansanti
Physics Department, Sapienza University of Rome and INFN sect. Roma1, P.le A. Moro 2, 00185 Rome, Italy
andrea.giansanti@roma1.infn.it

Keywords: Natively unfolded proteins; intrinsic disorder; unfoldome; diseasome; cell signalling, cancer.

Abstract: Intrinsically disordered proteins are fascinating the community of protein science since the last decade, at least. There is a well-established line of research that intends to reveal the crucial role played by intrinsically disordered proteins (IDPs) in the development of human diseases. The main argument is that IDPs are differentially more present in groups of disease-related proteins. In this note we compare the frequency of disorder in human proteins, both disease-related and not. The frequency of disorder is comparable in the two sub-groups of proteins. Disorder is widespread in human proteins, but it is not a specific pre-requisite of proteins involved in the development of cancer, cardiovascular diseases, diabetes and neurodegenerative diseases. A tendency of cancer-related proteins to be statistically more disordered than the rest of human proteins is confirmed.

1 INTRODUCTION

In this note we discuss the differential occurrence of intrinsically disordered proteins (IDPs) among disease-related proteins in the human proteome. On the basis of straightforward observations we raise some critical remarks about the often claimed strict association between the *unfoldome* (Uversky, 2010) and the human *diseasome* (Goh et al., 2007).

In the last decade a growing number of scientific publications have been devoted to IDPs; these are proteins lacking a well-defined three-dimensional tertiary structure in all or part of their polypeptide chain and existing as an ensemble of flexible conformations. It has been reported that IDPs fulfil important biological functions in the cell being involved in targeting, signalling and regulation of the cell cycle (Wright and Dyson, 1999; Dunker et al., 2001; Dunker et al., 2008; Tompa, 2010; Uversky, 2013; van der Lee et al., 2014). The absence of a tertiary structure suggests that these proteins do not use a lock-and-key mechanism to interact with their substrates, and a number of different mechanisms have been proposed, involving unfold-to-fold transitions and/or high specificity/low affinity interactions with target molecules (Dunker et al., 2001; Uversky, 2002; Tompa, 2010).

A growing interest is due also to the observation that IDPs might be essentially implied in cellular processes related to the development of human diseases. In particular, in a quite cited paper by Iakoucheva et al. of 2002, it was reported that cancer-related proteins are enriched in disorder. In another often quoted paper, it is reported that the fraction of IDPs is larger among proteins related to cardiovascular diseases (CVD) than in the SwissProt database, which, notably, contains proteins from all kingdoms of life (Cheng et al., 2006). Uversky et al. have clearly shown, in 2009, that large percentages of disease-related proteins are predicted as IDPs, though no comparison with the rest of proteins not explicitly related to diseases is made (see figure 4 in (Uversky et al., 2009)). Several other papers have linked IDPs with diseases (Midic et al., 2009; Uversky et al., 2014) and the D^2 concept (*disorder in disease*) has been introduced by Uversky in 2008 (Uversky et al., 2008). It has to be said

that many of the above mentioned papers adopt criteria to define a protein as IDP that overestimate the absolute number of IDPs in a data set, as we have previously shown (Deiana and Giansanti, 2013).

The strict association of protein disorder with complex pathologies, based on the claim of a systematic abundance of IDPs in disease-related proteins, has reverberated, since 2002, in a growing corpus of publications. From the whole of these studies one can be tempted to conclude that disorder plays indeed an essential role in the development of diseases with complex etiology, e.g. cancer, diabetes, cardiovascular and neurodegenerative syndromes. Looking back at the original studies and in most of the subsequent papers, the relatively high occurrence of IDPs in human disease-related proteins was compared with the occurrence of IDPs in sets of proteins from different organisms, which is an unfair reference. Therefore, it is incorrect to infer that, selecting human proteins annotated as disease-related, that corresponds to enriching in disordered proteins. As we show below, intrinsically disordered proteins have the same occurrence and distribution both among disease-related proteins and among the rest of the human proteome (HP) (figure 1). In a nutshell: there is no enrichment.

In this note, after a brief discussion of criteria to recognise a protein as disordered, we compare the percentage of IDPs occurring in human proteins annotated as disease-related and in the rest of HP. We show that the fraction of IDPs in the two sub-sets is similar and equally distributed over classes of disorder. Indeed, HP contains more IDPs than other organisms, but we did not found a specific enrichment among the disease-related proteins. An observation that could have been done earlier and that urge for a critical attitude in reading the copious literature about the connection between unfoldome and diseasome.

2 METHODS

The proteome of Homo Sapiens was downloaded from the SwissProt database, release of January 2011. According to the annotations in the database, the 20230 sequences in the HP were partitioned into two major subgroups: *disease-related* proteins (9110) and *the rest* (11120). Disease-related proteins were grouped then into subclasses, referring to four sets of diseases with a complex etiopathology, cancer (3427), cardiovascular diseases (CVD) (6645), diabetes (250) and neurodegenerative diseases (1380). Note that there are disease-related proteins that belong to more than one subclass. These subclasses were identified using the same set of keywords used by Uversky et al. in 2009.

Disordered residues in protein sequences were identified by DISOPRED2 (Ward et al., 2004), a well-known disorder predictor that has both good *sensitivity* and good *specificity*. Moreover, in a previous study, we found that it has also a good *selectivity*, an often overlooked index of performance that essentially controls the number of false positives (Deiana and Giansanti, 2013).

Usually protein sequences have been classified as intrinsically disordered if they contain at least one disordered segment longer than 30 amino acids (Oldfield et al., 2005; Le Gall et al., 2007). This criterion does not take into account the length of the protein and can induce an overestimation of the number of disordered proteins in a given set. Consider, for example, the case of a protein made by 1000 residues, with only one disordered segment of 30 residues. This protein would be classified as disordered notwithstanding the fact it could well contain structured regions in the rest of the sequence. A more sensible criterion is that of requiring, for a protein to be an IDP, to have at least 30% of disordered residues (Gsponer et al., 2008; Deiana and Giansanti, 2013).

In this study we adopt the intersection of both criteria and we classify a protein as an IDP if it has at least one disordered domain longer than 30 residues and at least 30% of residues that are predicted as disordered.

Uncertainties in the percentages of IDPs over classes of disorder (Figures 1 and 2) were estimated through bootstrap resampling and reported as standard errors (Efron and Tibshirani, 1993).

3 RESULTS

Following the two combined criteria mentioned above, we assessed the percentage of IDPs among disease-related and the rest of proteins in the human proteome. In table 1 the estimates based on the first and the second criterion, used either separatedly or in combination, are compared. Clearly, the second criterion is more conservative. Interestingly, if one adopts both criteria simultaneously one gets: 41% for the group of disease-related and 46% for the rest of proteins in the human proteome; these figures indicate that the second criterion is essential to reject, from both groups, spurious cases of proteins that are classified as disordered just because they comprise a disordered

segment longer than 30 residues, but nevertheless are not disordered in the major part of their sequences. The second criterion, that takes into account the fraction of disordered residues, reduces the incidence of IDPs in both groups by around 20%. In the following we adopt both criteria in combination. It is then interesting to estimate

Table 1: **Estimates of the percentage of IDPs in disease-related and the rest of the human proteome (HP).**
The first or the second criterion and the combination of the two (see text) are here considered. In the first criterion a protein is considered as intrinsically disordered if it has at least one disordered segment longer than 30 consecutive amino acids; in the second criterion a protein is disordered if more than 30% of its amino acids are predicted as disordered. Percentages are computed by normalising with respect to the number of proteins in each group.

GROUP	1st Crit.	2nd Crit.	Both criteria
	IDPs, %	IDPs, %	IDPs, %
Disease-related	0.66	0.43	0.41
Reast of HP	0.63	0.49	0.46

in disease-related human proteins and the rest of HP the probability that an IDP is disordered in at least a given percentage of its residues (see Figure 1), a kind of cumulative distribution. The distribution of IDPs over bins of percentages of disorder follow similar decreasing trends in both groups. Figure 1 clearly shows that, not only the incidence of IDPs in disease-related proteins is lower than in the rest of the human proteome globally, but also that it is lower in each bin of disorder, in detail.

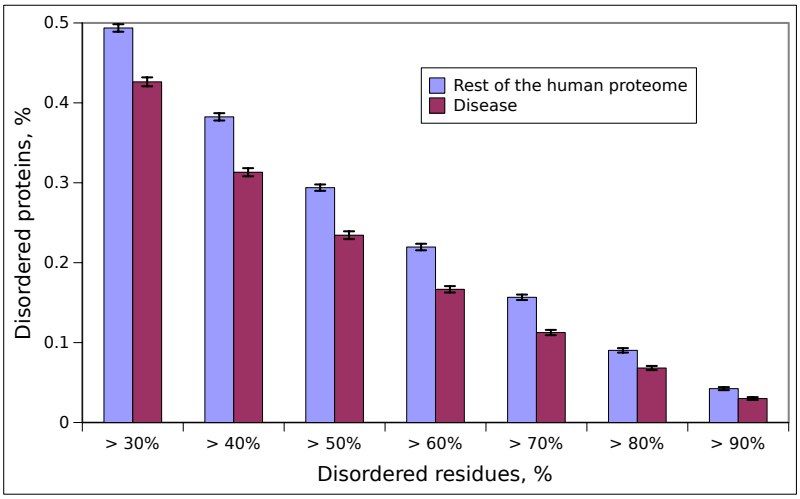


Figure 1: **Distribution of IDPs in disease-related proteins and in the rest of the human proteome.**
Disorder is here binned. The first bin groups IDPs with more than 30% of disordered residues; the second bin corresponds to proteins with more than 40% of disordered residues and so on. Clearly the second bin contains all the proteins in the first bin minus those proteins that have between 30% and 40% of predicted disordered residues. The frequency of occurrences is the number of IDPs in each bin divided, respectively, by the number of disease-related proteins and by the number of remaining proteins in the HP. The error bars represent standard errors (see section Methods) upon bootstrap resampling.

The human proteome has a large amount of IDPs, from the estimates in this note, it is close to 44%, but we cannot say that IDPs are more frequent among disease-related proteins. Disease-related proteins cannot be thought of as a group of proteins that are specifically enriched in disorder; they are disordered as much as the rest of the human proteome. Other interesting observations come from considering the cumulative distribution of disease-related IDPs over classes of disorder, separating various groups of diseases (see figure 2, where the occurrence of IDPs in the rest of HP is kept as a reference). The incidence of IDPs is different for different pathologies, though it tends to become uniform and comparable with that of the rest of HP in the classes with a large majority of disordered residues. Interestingly, in proteins related to cancer the occurrence of IDPs is higher than in the rest of the human proteome, in most classes of disorder. It becomes comparable with the rest of HP only from the bin of 70% on. IDPs related to neurodegenerative diseases follow next, but are less frequent than in cancer and in the rest of the HP, again becoming comparable in incidence only in the classes that are disordered for more than 80%. Proteins related to CVDs and diabetes appear to have less incidence of IDPs with respect to the other two subgroups and the rest of HP.

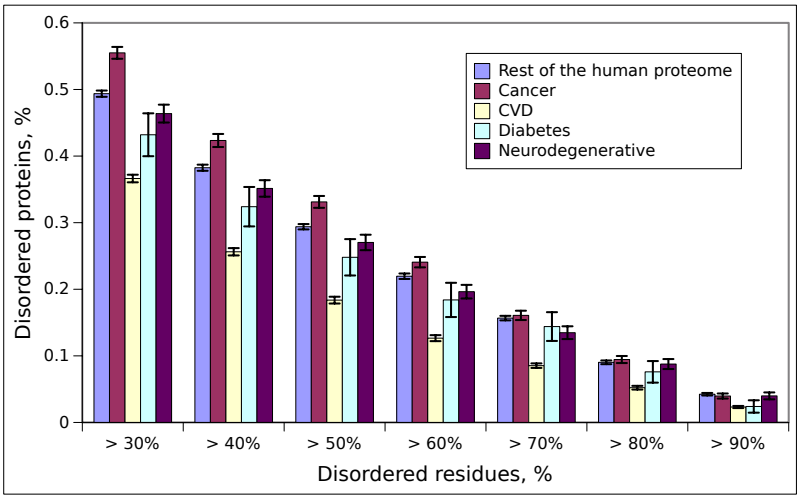


Figure 2: **Percent distribution of disease-related IDPs in classes of disorder for different groups of diseases**. Percentages of IDPs related to different groups of diseases are here binned over the same classes of disorder as in figure 1. The occurrence of IDPs in the rest of the human proteoome is reported as a reference.

4 DISCUSSION

Consistently with previous reports, we confirm that the fraction of IDPs is higher in cancer-related proteins with respect to the other groups (Iakoucheva et al., 2002; Uversky et al., 2009). Nevertheless it has to be noted that in those papers the comparison between the occurrence of IDPs in cancer-related human proteins is made with: signalling proteins, proteins in Swissprot from Eukaryots, and a non redundant non-homologous set of proteins from the PDB. No comparison is made, as we do here, with the rest of the HP. The remarkable enrichment of disorder in disease-related proteins, as shown in figure 1 of the paper by Iakoucheva et al. of 2002 and in figure 4 of the 2009 paper by Uversky et al., is due to the comparison made between disease-related human proteins and sets of proteins that mix up human proteins with proteins from other organisms that, as is well known, are biased towards less disorder. Moreover, in those original works the first criterion of IDPs classification was adopted, inducing an overestimate of the amount of disorder in all the considered instances, as we have discussed above. Notably, the biased estimates in (Iakoucheva et al., 2002), and (Uversky et al., 2009) propagated and somehow flawed many papers. Also recently we have found the same way of arguing, same type of graphs, in two interesting papers on the network of interactions governing cell death processes (Uversky, 2013; Peng et al., 2013).

As is well known one of the major players in the game of enforcing or escaping programmed cell death is the protein p53, which is inactivated in most cancers and has unstructured N- and C-terminal regions and a structured DNA-binding region. This protein is central to many discussions about the essential role of disorder in cancer, see, e.g., the articulate recent review by Uversky (Uversky et al., 2014). But it is worth noting that in this paper and in many that we have checked in the same thread no mention is made of the important observations published by Pajkos et al. in 2012. In that work (Pajkos et al., 2012) it is clearly suggested that there could be a problem of overestimation of disorder in previous works (Iakoucheva et al., 2002; Uversky et al., 2009). Moreover, Pajkos et al. observe, that in p53 cancer-related mutations and polymorphisms essentially occur in the structured regions and not in the disordered regions. This observation deserves due consideration in the current investigation of the role of disorder in the p53 machinery, but seems to have been overlooked.

It seems that the notion of intrinsic disorder in proteins still escapes a unitary physical-chemical definition. Possibly there are many variants of disorder and, nowadays after more than a decade of enthusiasm it would be necessary to make things clear. Even the term disorder might be out of place and should deserve clarification and exact definitions. We observe that a well-established line of research has a certain tendency at increasing the semantics in the field, as if IDPs were playing the role of the ”mysterious” dark matter of contemporary protein science (Uversky, 2010). From an operational point of view the concept of IDPs that is implemented in this line of research is, essentially, that of protein sequences enriched in disorder-promoting residues (i.e. charged and hydrophilic). Not much more than this. Looking at the future it must be said that very recently there have been two important meetings: see: `http://www.grc.org/programs.aspx?year=2014&program=`

idp and http://www.biophysics.org/2014dublin/Home/tabid/4526/Default.aspx) that point toward a more mature season of studies aimed at a detailed and specific definition of protein disorder. From a physical point of view even a globular crystallizable protein is, strictu sensu, a disordered, amorphous small piece of glassy material, i. e. lacking the periodicity of crystalline solids, which constitutes, somehow, the paradigm of ordered materials.

5 CONCLUSIONS

Intrinsically disordered proteins undoubtedly hold a peculiar fascination for the community of protein science. Under the enthusiast pressure of a group of devoted researchers the number of publications on the theme of IDPs is growing at a remarkable pace. Nevertheless, it must be noted that while on the one hand the number of papers on IDPs that are based on predictors is growing exponentially, on the other hand the number of circumstantiated experiments on specific protein molecules is growing perhaps logarithmically in time. Since doing experiments is more demanding than doing computations this last observation is almost trivial, one could say, but perhaps worth noting.

Our task in this note was just to warn that in dealing with protein disorder one has to face a deluge of publications that should be critically considered and not enthusiastically accepted, without checking, as the increasing evidence that disorder plays essential roles in fundamental biological processes related to the onset of complex diseases. We believe that, at the moment, more sober and circumstantiate approaches have to be implemented (e.g. (Babu et al., 2012)). A series of three papers had the merit, in 2007, to vigorously raise the problem of a systematic functional annotation of protein disorder (Xie et al., 2007b; Vucetic et al., 2007; Xie et al., 2007a). The two strictly related themes of functional annotation and evolutionary conservation of intrinsic disorder are masterly covered in the book by Peter Tompa (Tompa, 2010), see in particular chap. 11. In our opinion a correct evaluation of the incidence of disordered conserved domains in various functional protein classes is still to be accomplished and constitutes an important step toward a rational, evidence based, understanding of how the unfoldome is related to the human diseasome.

We conclude by saying that prevalence of disorder among disease-related proteins is not a rule. Indeed, cancer related proteins are, in our check, statistically more disordered than the rest of the HP. In our opinion, there is no need anymore to argue that protein disorder is important because of its high incidence among the proteins that are in many ways related to high-impact diseases like cancer, but simply because it is an interesting and pervasive mode of protein function, particularly in the human proteome. Disorder is common in the human proteome, but seems not a specific signature of diseases.

6 ACKNOWLEDGEMENTS

One of us (A.D.) has been enrolled, from 2007 to 2010, in cycle XXIII of the graduate school of Biophysics of Sapienza University of Rome, at the time part of the CISB. Both of us warmly tank prof. Alfredo Colosimo for the stimulating atmosphere and the support during those years; the study of intrinsic disorder in proteins has been the main theme of Antonio Deiana's doctoral thesis.

REFERENCES

Babu, M. M., Kriwacki, R. W., and Pappu, R. V. (2012). Structural biology. versatility from protein disorder. *Science*, 337(6101):1460–1.

Cheng, Y., LeGall, T., Oldfield, C. J., Dunker, A. K., and Uversky, V. N. (2006). Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry*, 45(35):10448–60.

Deiana, A. and Giansanti, A. (2013). Tuning the precision of predictors to reduce overestimation of protein disorder over large datasets. *J Bioinform Comput Biol*, 11(2):1250023.

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001). Intrinsically disordered protein. *J Mol Graph Model*, 19(1):26–59.

Dunker, A. K., Oldfield, C. J., Meng, J., Romero, P., Yang, J. Y., Chen, J. W., Vacic, V., Obradovic, Z., and Uversky, V. N. (2008). The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics*, 9 Suppl 2:S1.

Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*, volume 57. Chapman & Hall, New York.

Goh, K.-I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., and Barabási, A.-L. (2007). The human disease network. *Proc Natl Acad Sci U S A*, 104(21):8685–90.

Gsponer, J., Futschik, M. E., Teichmann, S. A., and Babu, M. M. (2008). Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science*, 322(5906):1365–8.

Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z., and Dunker, A. K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol*, 323(3):573–84.

Le Gall, T., Romero, P. R., Cortese, M. S., Uversky, V. N., and Dunker, A. K. (2007). Intrinsic disorder in the protein data bank. *J Biomol Struct Dyn*, 24(4):325–42.

Midic, U., Oldfield, C. J., Dunker, A. K., Obradovic, Z., and Uversky, V. N. (2009). Protein disorder in the human diseasome: unfoldomics of human genetic diseases. *BMC Genomics*, 10 Suppl 1:S12.

Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N., and Dunker, A. K. (2005). Comparing and combining predictors of mostly disordered proteins. *Biochemistry*, 44(6):1989–2000.

Pajkos, M., Mészáros, B., Simon, I., and Dosztányi, Z. (2012). Is there a biological cost of protein disorder? analysis of cancer-associated mutations. *Mol Biosyst*, 8(1):296–307.

Peng, Z., Xue, B., Kurgan, L., and Uversky, V. N. (2013). Resilience of death: intrinsic disorder in proteins involved in the programmed cell death. *Cell Death Differ*, 20(9):1257–67.

Tompa, P. (2010). *Structure and function of intrinscially disordered proteins*. Chapman and Hall/CRC Press, Boca Raton.

Uversky, V. N. (2002). What does it mean to be natively unfolded? *Eur J Biochem*, 269(1):2–12.

Uversky, V. N. (2010). The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. *J Biomed Biotechnol*, 2010:568068.

Uversky, V. N. (2013). Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta*, 1834(5):932–51.

Uversky, V. N., Davé, V., Iakoucheva, L. M., Malaney, P., Metallo, S. J., Pathak, R. R., and Joerger, A. C. (2014). Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem Rev*, 114(13):6844–79.

Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: introducing the d2 concept. *Annu Rev Biophys*, 37:215–46.

Uversky, V. N., Oldfield, C. J., Midic, U., Xie, H., Xue, B., Vucetic, S., Iakoucheva, L. M., Obradovic, Z., and Dunker, A. K. (2009). Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics*, 10 Suppl 1:S7.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E., and Babu, M. M. (2014). Classification of intrinsically disordered regions and proteins. *Chem Rev*, 114(13):6589–631.

Vucetic, S., Xie, H., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Obradovic, Z., and Uversky, V. N. (2007). Functional anthology of intrinsic disorder. 2. cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *J Proteome Res*, 6(5):1899–916.

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol*, 337(3):635–45.

Wright, P. E. and Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol*, 293(2):321–31.

Xie, H., Vucetic, S., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Obradovic, Z., and Uversky, V. N. (2007a). Functional anthology of intrinsic disorder. 3. ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J Proteome Res*, 6(5):1917–32.

Xie, H., Vucetic, S., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Uversky, V. N., and Obradovic, Z. (2007b). Functional anthology of intrinsic disorder. 1. biological processes and functions of proteins with long disordered regions. *J Proteome Res*, 6(5):1882–98.

Current Trends in the Mathematical Modeling of the Cellular Metabolism

Hermann-Georg Holzhütter

University Medicine Berlin (Charité) - Institute of Biochemistry - Charitéplatz 1, 10117 Berlin
Mail: hergo@charite.de

Introduction

Mathematical modeling of the cellular metabolism has a long tradition that started in the early 1970s with detailed kinetic models of central metabolic pathways of cardiomyocytes, yeast cell and erythrocytes. In the post-genomic era, computational methods become increasingly important for putting experimental information from various layers of molecular organization (mRNA- and protein profiles, metabolite patterns, enzyme-kinetic data) into consistent network models of cellular metabolism (Knupfer et al., 2013). The results of novel high-throughput technologies as, for example, RNA- and DNA Microarrays, Mass-Spectroscopy (MS) or Reverse Phased Protein Microarrays are being stored in exponentially growing data bases of cellular macromolecules and low-molecular weight compounds providing 'molecular pictures' of cells and tissues under different physiological settings. The traditional concept of kinetic modeling (KM) based on a detailed mathematical description of all elementary reactions involved in a reaction network appears to be not practicable for the dynamic study of genome-wide metabolic networks as the regulatory properties of the majority of enzymes and membrane transporters, in particular those operating at the level of gene regulation and signal transduction, are yet not well-characterized enough. Therefore, in order to distill knowledge from 'omics' data sets, novel computational methods have been proposed that collectively are referred to as topology-based models. They have in common that the topology of the metabolic network given in terms of metabolites and biochemical reactions often serves as the only experimental input whereas kinetic details of the individual reactions are disregarded. The most popular of these novel methods is constraint-based modeling (CBM). This method relies on the assumption that the regulation of metabolism has been shaped during natural evolution to enable an optimal adaptation of enzymatic reactions rates (usually called fluxes) to varying physiological requirements as cell growth rate or defense against harmful agents. Under this premise, the determination of metabolic fluxes reduces to the optimization of an objective function relating the metabolic fluxes to the considered physiological task(s). In the following I review the biological insights that my group has gained into the architecture and regulation of the liver metabolism during the past ten years based on applications of CBM and KM, thereby critically assessing the pros and cons of either methods.

Basic model equations

Both CBM and KM eventually rest on a description of metabolic networks that assumes spatial homogeneity of metabolite concentrations within a cellular compartment and relates the time-dependent variation of metabolite concentrations to the rates of enzymes and transporters according to the rules of chemical kinetics (Heinrich and Schuster, 1998)

$$\frac{d[M_i]}{dt} = \sum_{i=1}^{n} S_{ij}\, v_j \quad (i=1,\ldots,m) \qquad (1)$$

The left-hand side symbolizes the time-dependent concentration change of metabolite (i), S_{ij} is the so-called stoichiometric matrix composed of the (integer) stoichiometric coefficients with which metabolites (i) are produced or consumed in reactions (j) and v_j is the reaction rate. The complete set of reaction is commonly referred to a *flux distribution*. Stationary states are defined through the condition $d[M_i]/dt = 0$

The reaction rates depend on the concentration of the reactants (substrates, products) as well as on other types of regulation (allosteric effectors, binding of metal ions, reversible chemical modifications, dissociation into subunits etc.) rendering the enzymatic rate laws complicated non-linear multi-parametric expressions. The most comprehensive repository of enzymatic rate laws in the public is the data base BRENDA (Scheer et al., 2011). In rare cases where available information allows to take into account the intra-cellular spatial

distribution of metabolites and enzymes the model equations (1) have to be replaced by systems of partial differential equations.

Topology-based modeling of liver metabolism

The liver metabolism plays an important role for the homeostasis of numerous plasma metabolites (e.g. glucose, fatty acids, lipoproteins, amino acids), the formation of the bile and the detoxification of xenobiotic (drugs, toxins) and internal waste products (e.g. bilirubin, ammonia). This large pattern of metabolic transformations is almost exclusively brought about by hepatocytes. Therefore, understanding the regulation of the hepatocyte's metabolism is of great importance for giving scientifically sound recommendations for a liver-friendly diet, for designing novel pharmacological therapies supporting or re-directing the liver metabolism in certain systemic diseases (inflammation, diabetes) and for preventing the development of liver diseases (alcoholic and non-alcoholic fatty liver and related inflammatory processes).
Starting point for metabolic modeling is the reconstruction of the network under consideration. Whereas the enzymatic reactions and membrane transport processes involved in central metabolic pathways as glycolysis, gluconeogenesis, the pentose phosphate pathway, the mitochondrial tricarbonic acid (TCA) cycle or the urea cycle are known since decades, an (almost) complete reconstruction of the metabolic network of hepatocytes was reported only recently. In (Jerby et al., 2010), the network reconstruction was based on an algorithm that generates a tissue-specific model from the generic human model by integrating a variety of tissue-specific molecular data sources, including literature-based knowledge and 'omics' data. HEPATONET1, the reconstruction in (Gille et al., 2010) was based on a thorough study of biochemical literature complemented by inclusion of available liver-specific gene expression data. Great value was set upon testing the consistency of the network (Hoffmann et al., 2007) and its functional capacity to carry out a large array of known liver-specific metabolic tasks. This process of so-called network 'curation' can be comfortably performed by using constraint-based modeling (CBM). The core of this method, also termed flux-balance analysis (FBA), is to construct a stationary flux distribution that satisfies the steady-state condition and optimizes (i.e. either maximizes or minimizes) a user-defined flux objective function $\Phi(v_1, v_2, ..., v_n)$ depending on a set of physiologically relevant network fluxes (often called target fluxes) (Orth et al., 2010). In its simplest form the mathematical formulation of a constraint-based optimization problem reads

$$\Phi(v_1, v_2, ..., v_n) \rightarrow \text{MIN(MAX)!}$$

$$\frac{d[M_i]}{dt} = \sum_{i=1}^{n} S_{ij} v_j = 0 \qquad (2)$$

$$-lb_j \leq v_j \leq ub_j$$

i.e. the condition of a metabolic steady-state (vanishing time variations of the metabolite concentrations) and the prescription of lower and upper bounds (lb_j, ub_j) for the fluxes constrain the optimization of the flux objective. A further important constraint arising from basic thermodynamics (but often disregarded in CBM applications) is the condition that the direction of fluxes has to be in concordance with the sign of the change of the Gibbs free energy (i.e. $v_j > 0$ if $\Delta G < 0$) (Hoppe et al., 2007).
For example, testing whether the reconstructed network allows the production of bile at given external medium of the hepatocyte, the fluxes producing individual bile components (as cholesterol, phospholipids and bile acids) have to be put on fixed numerical values ($lb_j = ub_j = L_j$ in (2)) according to the relative proportion with which they occur in the bile. Solution of the optimization problem (2) selects the "best" flux distribution among the commonly infinitely large set of alternative flux distributions (see the simple example in Fig. 1). Different flux objectives have been proposed to determine physiologically feasible flux distributions (systematically investigated in (Schuetz et al., 2007)). We commonly use the minimization of internal fluxes (Holzhütter, 2004) as flux objective and seek solutions that are thermodynamically feasible.
Besides being a useful tool for assessing the consistency and functional capacity of a metabolic network, constraint-based optimization can also be used to identify essential reactions and to estimate the maximal demand/yield of substrates/products (see Fig. 1).

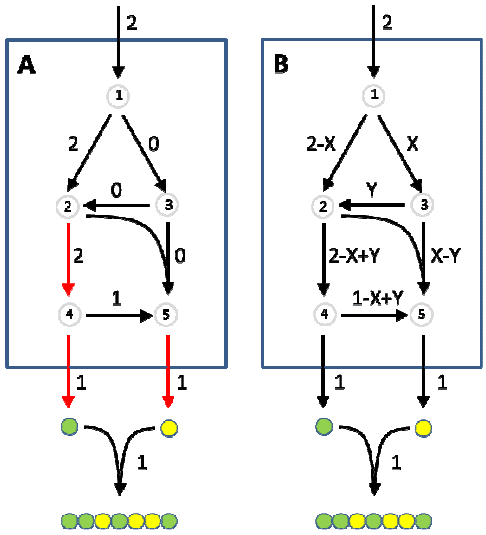


Fig. 1 *Computation of metabolic fluxes by means of constraint-based modeling (see (2))*
The simplistic network converts one input metabolite to two output metabolites serving as building blocks for a macromolecule (= biomass production). The task consists in determining all metabolic routes producing one unit of biomass.
A Optimal flux mode rendering the sum of fluxes minimal
B There exists an infinite set of possible flux distributions with 0<=X<=2, Y>=X producing one unit of biomass.
The red arrows symbolize essential reactions.

Essential reactions are defined as reactions which upon elimination from the network (= synthetic knock-out) render at least one physiologically relevant biochemical reaction impossible to carry a non-zero flux. Essential reactions are convenient targets for the development of drugs impairing the metabolism of pathogens (see, for example, (Huthmacher et al., 2010; Bazzani et al., 2012)). Moreover, a deficiency of enzymes catalyzing essential reactions is frequently associated with the occurrence of metabolic disorders: 80 reactions out of 207 essential reactions identified in HEPATONET1 are reported to be associated with metabolic diseases.

Assessment of the maximal yield of a specific cellular target metabolite (e.g. ATP, triglyceride, cholesterol) obtainable per unit utilized substrates is another useful application of CBM. For example, using HEPATONET1 to quantify the amount of utilized oxygen (roughly proportional to the amount of ATP produced in oxidative phosphorylation) required to convert (toxic) ammonia in one unit of he alternative nitrogen sinks glutamine, asparagine or urea revealed urea formation to represent the by far most ATP demanding process. This finding may explain the spatial zonation of hepatic ammonia detoxification: Hepatocytes in the periportal region of the liver sinusoids sensing the highest capillary oxygen levels convert ammonia into urea whereas hepatocytes in the pericentral region (see explanation in Fig. 2B) exposed to the lowest oxygen concentration possess a high capacity of the enzyme glutamine synthetase. Another useful application of CBM-based yield calculations was carried out in (Karlstädt et al., 2012). Taking the synthesis of ATP, the lipids ceramide and cardiolipin and further important phospholipids as examples, it was analyzed how changes in the relative supply of glucose, lactate, fatty acids and ketone bodies may influence the efficiency of these essential processes. The results of such yield calculations maybe particularly useful for the purposeful supplementation of nutrients which otherwise restrain the optimal accomplishment of central metabolic functions

The striking advantage of CBM is its simplicity: It requires only knowledge of the stoichiometry of the network and thus can be immediately applied to whole-cell network reconstructions without caring for the kinetic and regulatory properties of enzymes and transporters. On the other hand, the predicted flux distributions are hypothetical and have to be considered with care because they are essentially dictated by the subjective choice of the flux objective. As the metabolism enables a large number of diverse cellular functions at the same time, the assumption of a single flux objective does not seem realistic. For example, liver cells (hepatocytes) contribute to the homeostasis of the plasma glucose concentration by either producing or utilizing glucose. Thus, the glucose exchange flux (which can be either positive or negative) is one of the outmost important metabolic objectives of hepatocytes. The magnitude of this flux is controlled by a large array of effectors, e.g. the plasma level of the hormones insulin, glucagon, epinephrine and somatostatin (to name only the most important ones) or the plasma level of fatty acids which are able to reduce the rate of glucose uptake (the so-called Randle effect). All these regulatory factors are not part of a CSB model. So how can a CSB modeler formulate a feasible constraint for the glucose exchange flux or fix its value? Moreover, glucose exchange is only one important task of hepatocytes besides about 80 other equally important functions reported for the human liver that, in principal, have to be included into a reliable CSB model (Gille et al., 2010). Specifying realistic values for the output fluxes related to these functions meets the same difficulties as outlined before for the specification of the glucose exchange flux. Whereas the

formulation of a realistic multi-objective function appears to be impossible for complex eukaryotic cells like hepatocytes, the situation looks slightly better for simple prokaryotes like E. coli. Comparing 44 experimentally determined in vivo MFA datasets for E. coli to predicted flux distributions based on a large number of putative flux objectives and linear combinations of these objectives, the best concordance was achieved by a combination of three flux objectives: maximizing biomass production, maximizing ATP generation, and minimizing reaction fluxes across the network (Schuetz et al., 2012). But still the relative proportions among these three objectives may vary between different cells putting the predictive value of such multi-objective function into question. Another serious limitation of CBM is the impracticality to assess metabolite concentrations. As a surrogate solution for predicting changes of the metabolite profile elicited by network perturbations, Patil and Nielsen proposed the concept of so-called reporter metabolites being those metabolites around which the largest expression changes of enzymes occur (Patil and Nielsen, 2005). Mapping transcriptomics data obtained from patients with non-alcoholic fatty liver disease (NAFLD) to a metabolic network of human hepatocytes, Mardinoglu et al. identified chondroitin and heparan sulphates as suitable reporter metabolites for the presence of a non-alcoholic steatohepatitis (NASH) and proposed three enzymes of the serine pathway as potential therapeutic targets for the treatment of NASH (Mardinoglu et al., 2014).

Kinetic modeling of liver metabolism

Owing to the need of detailed biochemical information on the reaction kinetics and enzyme-regulatory properties for the individual reactions and transport processes, kinetic models of liver metabolism have been hitherto confined to small metabolic sub-networks with glycolysis being the most intensively modelled pathway so far. Biological and medical questions addressed by these model pertain the interdependence of gluconeogenesis and fatty-acid catabolism (Chalhoub et al., 2007), the impairment of glucose production in von Gierke's and Hers' diseases (Beard and Qian, 2005), the role of the pyruvate metabolism (Stucki and Urbanczik, 2005) or the implication of metabolic zonation (Ohno et al., 2008). Each of these models focuses on a small number of reactions pertinent to the metabolic function of interest embedded in a customized representation of the principal pathways of central metabolism. Recently, our group developed a more complex kinetic model of the glucose metabolism of hepatocytes comprising the synthesis and degradation of glycogen, glycolysis and gluconeogenesis (König et al., 2012; König and Holzhütter, 2012). A special feature of this model is the inclusion of hormonal regulation of the liver metabolism by insulin and glucagon. This was accomplished by taking into account the cAMP-dependent reversible phosphorylation of key regulatory enzymes (e.g. glycogen synthetase, glycogen phosphorylase, pyruvate kinase). The fraction of phosphorylated enzyme is regulated by an intricate signaling network centered on the cardinal kinases protein kinase A (PKA) and protein kinase B (AKT). Instead of explicit modelling of this signaling network, we used an empirical Hill-type signal function to relate the fraction of phosphorylated enzymes to the plasma level of the hormones. Model simulations revealed that in the liver of subjects suffering from diabetes type 2 the set point of glucose exchange (at which hepatic glucose uptake equals glucose release) is shifted from the normal value of about 6 mM to a significantly higher value of about 8 mM, thus rendering the liver for plasma glucose levels below 8 mM to a permanent glucose producer. Intriguingly, model simulations suggested that dropping the plasma level of glucose to the normal value by means of a rigorous insulin therapy increases the risk of hyperglycemic episodes (too low plasma glucose level endangering especially the energy metabolism of the brain).

One of the central aims of molecular systems biology is to bridge the different scales inherent the hierarchical organization of higher multi-cellular organisms in molecules, cells, tissues and organs. A successful multi-scale modeling approach integrating the electrophysiology of the cardiomyocyte into a finite-element tissue model of muscle deformation is the heart model (Fink et al., 2011) developed in the cardiac 'Physiome' project, a research initiative of the International Union of Physiological Sciences (IUPS) attempting to provide a comprehensive framework for modelling the human body using computational methods. For the liver, a concept for a multi-scale mathematical model has been proposed (Holzhütter et al., 2012) that considers the organ as a parallel connection of a huge number (more than a million) of liver lobules. Each lobule is composed in a canonical way by a hexagonal arrangement of liver cells lining the liver capillaries (sinusoids). Following to this modeling concept, we have started to build a multi-scale metabolic model of the liver metabolism (yet unpublished work, see Fig. 2). Such a multi-scale model allows to simulate the impact of alterations in the hepatic blood perfusion (e.g. due to the aging of the organ or in the presence of liver fibrosis and cirrhosis) on the metabolic capacity of the organ. A multi-scale tissue model allows to take into

account both the heterogeneity of metabolite and hormone concentrations along the sinusoid as well as the differential endowment of hepatocytes with metabolic pathways. As many hormones are efficiently cleared from the plasma during their passage through the liver capillaries, their plasma concentration and thus their signaling effects in the hepatocytes largely varies between the arterial and venous part of the vessel. This is accompanied by location-dependent variations in the expression level of key regulatory enzymes of important pathways like glycolysis, gluconeogenesis and urea synthesis different expression levels ('metabolic zonation'). These heterogeneities may result in quite distinct metabolic activities in different cell populations. For example, the accumulation of triglycerides in the fatty liver has been hypothesized to increase the volume of hepatocytes and this way to compress the diameter of sinusoids. This would create a vicious cycle in that the reduced blood flow (less oxygen) promotes the accumulation of triglycerides (Schleicher et al., 2014). We intend to use the multi-scale model to check whether such a mechanism may indeed explain the observed regional heterogeneity of lipid deposits in the liver.

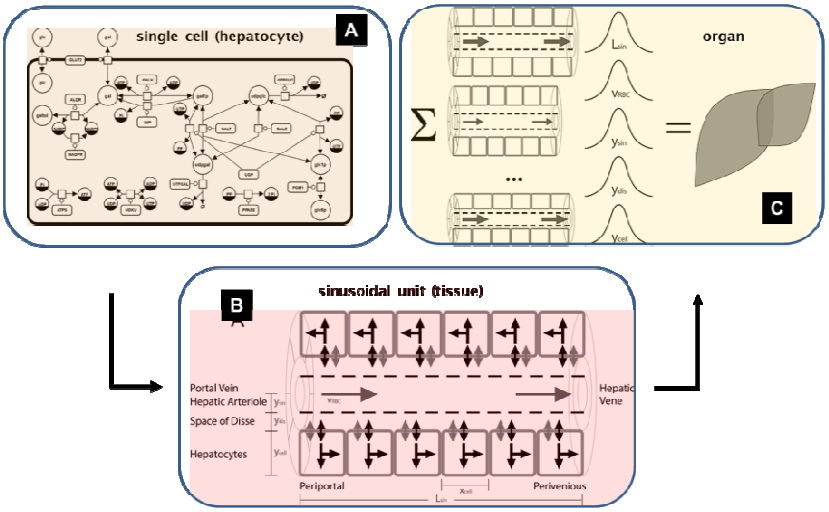


Fig. 2 *Schematic representation of a multi-scale model of liver metabolism.*
At the cellular level (hepatocyte), the metabolism is modeled by kinetic models taking into account the regulation of enzymes and transporters by allosteric effectors, reversible phosphorylation and gene expression. At the tissue level of a single sinusoidal unit the exchange of metabolites and hormones between the blood, the space separating sinusoids from hepatocytes (space of Disse) and the cells is described by reaction-diffusion equations. At given plasma profile, the metabolic input-output relationship is computed for a large set of sinusoidal units differing in geometric parameters (e.g. thickness and length of vessels) and blood flow velocities sampled from experimentally determined distribution functions. Finally, the contribution of the sinusoidal units to the uptake/release of blood compounds is integrated to yield the input-output behavior of the organ.

Conclusion and outlook

Currently we are witnessing a renewed scientific interest in the study of the cellular metabolism, particularly in the fields of cancer research and neurobiology. Most of the contemporary work in metabolic research focuses on the regulation of gene expression. As the knowledge of the mechanistic details of gene regulation is rudimentary yet, topology base modeling techniques are usually exploited to infer changes in the cellular metabolism from changes in the expression level of enzymes. However, it is getting more and more obvious that changes in the abundance of enzymes and membrane transporters is only one and possibly not even the most important mode of metabolic regulation. Many key regulatory enzymes do not exhibit significant changes in their abundance despite large flux changes (ter Kuile and Westerhoff, 2001; Vogt et al., 2002). Thus, to understand metabolic regulation in a detail that enables a targeted intervention into metabolic processes with the aim to prevent or cure diseases, it appears to be indispensable to expand the concept of kinetic models to larger parts of the cellular metabolic network including different pathways resident in different organelles. The often said excuse that there are not enough kinetic data to build large-scale kinetic models of the cellular metabolism is simply not true. Several generations of brave biochemists have compiled a huge mountain of kinetic information on most enzymes of the intermediary metabolism of different cell

types. These data await their usage in mechanistic models of the cellular metabolism. Using kinetic models to correlate observed changes in the gene expression of enzymes with changes of maximal enzyme activities may be an effective means to elucidate strategies of cellular adaptation.
Another important aspect of future modeling projects concerns the inclusion of tissue perfusion. The functionality of a cell is largely determined by its access to nutrients and oxygen in the blood. Hypoxia and ischemia are the main causes for cellular dysfunctions and cell death. Thus, multi-scale models taking into account the (to a large part metabolic) regional and central regulation of the blood flow are essential to better understand the phenomenon of progressive tissue demolition, for example, in neurodegenerative diseases or after an acute infarction (heart, brain).
Finally: It is my personal conviction, that the future reputation and acceptance of mathematical models by experimental scientists, physicians will largely depend on the rigorousness with which these models truly take into biology and are clearly purpose oriented.

References

Bazzani S, Hoppe A, Holzhütter HG (2012) Network-based assessment of the selectivity of metabolic drug targets in Plasmodium falciparum with respect to human liver metabolism. BMC systems biology 6:118.

Beard DA, Qian H (2005) Thermodynamic-based computational profiling of cellular regulatory control in hepatocyte metabolism. American journal of physiology Endocrinology and metabolism 288:E633-644.

Chalhoub E, Hanson RW, Belovich JM (2007) A computer model of gluconeogenesis and lipid metabolism in the perfused liver. American journal of physiology Endocrinology and metabolism 293:E1676-1686.

Fink M, Niederer SA, Cherry EM, Fenton FH, Koivumaki JT, Seemann G, Thul R, Zhang H, Sachse FB, Beard D, Crampin EJ, Smith NP (2011) Cardiac cell modelling: observations from the heart of the cardiac physiome project. Progress in biophysics and molecular biology 104:2-21.

Gille C, Bolling C, Hoppe A, Bulik S, Hoffmann S, Hübner K, Karlstädt A, Ganeshan R, König M, Rother K, Weidlich M, Behre J, Holzhütter HG (2010) HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. Molecular systems biology 6:411.

Heinrich R, Schuster S (1998) The modelling of metabolic systems. Structure, control and optimality. Bio Systems 47:61-77.

Hoffmann S, Hoppe A, Holzhütter HG (2007) Pruning genome-scale metabolic models to consistent ad functionem networks. Genome informatics International Conference on Genome Informatics 18:308-319.

Holzhütter HG (2004) The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. European journal of biochemistry / FEBS 271:2905-2922.

Holzhütter HG, Drasdo D, Preusser T, Lippert J, Henney AM (2012) The virtual liver: a multidisciplinary, multilevel challenge for systems biology. Wiley interdisciplinary reviews Systems biology and medicine 4:221-235.

Hoppe A, Hoffmann S, Holzhütter HG (2007) Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. BMC systems biology 1:23.

Huthmacher C, Hoppe A, Bulik S, Holzhütter HG (2010) Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis. BMC systems biology 4:120.

Jerby L, Shlomi T, Ruppin E (2010) Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. Molecular systems biology 6:401.

Karlstädt A, Fliegner D, Kararigas G, Ruderisch HS, Regitz-Zagrosek V, Holzhütter HG (2012) CardioNet: a human metabolic network suited for the study of cardiomyocyte metabolism. BMC systems biology 6:114.

Knupfer C, Beckstein C, Dittrich P, Le Novere N (2013) Structure, function, and behaviour of computational models in systems biology. BMC systems biology 7:43.

König M, Holzhütter HG (2012) Kinetic modeling of human hepatic glucose metabolism in type 2 diabetes mellitus predicts higher risk of hypoglycemic events in rigorous insulin therapy. The Journal of biological chemistry 287:36978-36989.

König M, Bulik S, Holzhütter HG (2012) Quantifying the contribution of the liver to glucose homeostasis: a detailed kinetic model of human hepatic glucose metabolism. PLoS computational biology 8:e1002577.

Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J (2014) Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nature communications 5:3083.

Ohno H, Naito Y, Nakajima H, Tomita M (2008) Construction of a biological tissue model based on a single-cell model: a computer simulation of metabolic heterogeneity in the liver lobule. Artificial life 14:3-28.

Orth JD, Thiele I, Palsson BO (2010) What is flux balance analysis? Nature biotechnology 28:245-248.

Patil KR, Nielsen J (2005) Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proceedings of the National Academy of Sciences of the United States of America 102:2685-2689.

Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, Sohngen C, Stelzer M, Thiele J, Schomburg D (2011) BRENDA, the enzyme information system in 2011. Nucleic acids research 39:D670-676.

Schleicher J, Guthke R, Dahmen U, Dirsch O, Holzhuetter HG, Schuster S (2014) A theoretical study of lipid accumulation in the liver-implications for nonalcoholic fatty liver disease. Biochimica et biophysica acta 1841:62-69.

Schuetz R, Kuepfer L, Sauer U (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. Molecular systems biology 3:119.

Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U (2012) Multidimensional optimality of microbial metabolism. Science 336:601-604.

Stucki JW, Urbanczik R (2005) Pyruvate metabolism in rat liver mitochondria. What is optimized at steady state? The FEBS journal 272:6244-6253.

ter Kuile BH, Westerhoff HV (2001) Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. FEBS letters 500:169-171.

Vogt AM, Poolman M, Ackermann C, Yildiz M, Schoels W, Fell DA, Kubler W (2002) Regulation of glycolytic flux in ischemic preconditioning. A study employing metabolic control analysis. The Journal of biological chemistry 277:24411-24419.

DECREASED PARIETAL BUT INCREASED FRONTAL AMPLITUDE OF LOW-FREQUENCY FLUCTUATIONS OF BLOOD OXYGENATION COEXISTS WITH THE WIDESPREAD LOSS OF FUNCTIONAL CONNECTIVITY IN ALZHEIMER'S DISEASE

Mauro DiNuzzo, Daniele Mascali, Bruno Maraviglia, Laura Serra, Marco Bozzali, Federico Giove

Museo Storico della Fisica e Centro di Studi e Ricerche "Enrico Fermi", Rome, Italy
Dipartimento di Fisica, Università Sapienza, Rome, Italy
Fondazione Santa Lucia IRCCS, Rome, Italy
mauro.dinuzzo@neuroenergetics.org, daniele.mascali@uniroma1.it, bruno.maraviglia@roma1.infn.it, l.serra@hsantalucia.it, m.bozzali@hsantalucia.it, federico.giove@roma1.infn.it

Keywords: Brain fMRI; Resting-State; Alzheimer's disease.

Abstract: Alzheimer's disease (AD) is the most common form of age-related neurological disorder. AD patients suffer of cognitive and intellectual deficits that have been related to several biomarkers, including molecular aberrations as well as vascular and metabolic dysfunctions. On the side of functional brain dynamics, both electroencephalographic (EEG) and functional magnetic resonance imaging (fMRI) studies have revealed that cortical rhythms exhibit a loss of spatial coherence among different brain areas. Furthermore, a common finding in AD is that the power spectrum of spontaneous EEG exhibits a decrease of high and an increase of low frequencies. In this paper, we sought to examine whether there is a corresponding change in the amplitude low-frequency fluctuations (ALFF) of the blood-oxygenation level-dependent (BOLD) fMRI signal during resting-state. We have found that in AD patients ALFF decreases in the posterior cingulate cortex (PCC) and increases in prefrontal cortex, two critical constituent regions of the brain default-mode network (DMN). The changes in ALFF occur on top of the well-known widespread reduction of DMN spatial extension and mean seed-based functional connectivity (FC) as well as gray-matter volume. Overall, our results indicate that AD distinctly affects ALFF and FC. We discuss the relevance of this finding to the potential role of non-neuronal mechanisms in the pathophysiology of AD.

1 INTRODUCTION

Alzheimer's disease (AD) is the most frequent cause of dementia among elderly population (Takashima, 2009). AD is characterized by diffuse neuronal loss, accumulation of β-amyloid and tau proteins and development of senile plaques and neurofibrillary tangles in the cerebral cortex (Selkoe, 1994). Although these pathological markers are found in all cortical regions, the association areas are those most vulnerable to neurodegeneration, while primary motor and sensory areas are relatively preserved (Braak et al., 1993), consistent with a major impairment of cognitive functions.

Historically, one of the first hallmark of AD has been provided by the study of brain oscillations. It was Hans Berger (1873-1941), the recorder of the first electroencephalogram, to discover the AD-related pathological abnormalities in the rhythmic cerebral electrical activity. Thereafter, many studies confirmed that the electroencephalography (EEG) signature in AD is an increase of the theta (4-8 Hz) and delta (<4 Hz) activities and a decrease of alpha (8-12 Hz) and beta (12-30 Hz) activities, as well as a reduced coherence of the alpha and beta bands between distinct cortical regions (reviewed by Jeong, 2004). Specifically, the cognitive decline appears to be accompanied by higher low-frequency power in frontal-central cortical areas and lower high-frequency power in parietal-occipital cortical areas (Claus et al., 1998). These changes have been correlated with the degree of dementia (Cook and Leuchter, 1996) and also with its progression from amnestic patients suffering of mild cognitive impairment (MCI) (Grunwald et al., 2002), a preclinical prodrome of AD.

Support to the notion that the cognitive losses underlying AD arise from cortical disconnections comes from studies of functional magnetic resonance imaging (fMRI). In particular, analysis of resting-state low-frequency (<0.1 Hz) oscillations of the blood-oxygenation level-dependent (BOLD) fMRI signal reveals a progressive disruption of functional connectivity (FC) in the brain of MCI and AD patients compared with healthy controls. The loss of FC occurs predominantly in regions identifying the default-mode network (DMN) of the brain, including posterior cingulate cortex (PCC), prefrontal cortex (PFC) and inferior temporal/parietal cortex (Greicius et al., 2003b, He et al., 2007, Zhang et al., 2010, Gili et al., 2011). Whether the reduction of FC in dementia is accompanied by a variation of the amplitude of low frequency fluctuations (ALFF) within the DMN areas remains poorly investigated. Recently, it has been shown that MCI and AD patients exhibit increased ALFF in lateral temporal and superior frontal-parietal regions as well as decreased ALFF in medial parietal regions (including PCC) compared with healthy controls (Wang et al., 2011). In the present paper, we sought to compare the changes in FC and ALFF occurring in the cerebral cortex of AD patients relative to healthy subjects by using a region-of-interest (ROI) based approach.

2 MATERIALS AND METHODS

2.1 SUBJECTS

Seventeen patients diagnosed with probable AD by NINCDS-ADRDA consensus criteria (McKhann et al., 1984) and twenty healthy elderly controls (HC), with no evidence of cognitive deficits on neuropsychological testing, participated in this study. Major systemic, psychiatric and other neurological illnesses were carefully investigated and excluded in all studied subjects. In order to minimize the risk of concomitant cerebrovascular disease, subjects were excluded if they had either two or more hyperintense lesions with a diameter of 10 mm, or more than eight hyperintense lesions with a diameter between 5 to 9 mm on dual-echo MR images. Patients and controls were matched for gender and age. As expected, AD and HC scored differently in Mini-Mental State Examination (p<0.001). The study was approved by ethics committee of Santa Lucia Foundation; all subjects gave written consent before study.

2.2 DATA ACQUISITION

All subjects underwent a resting state fMRI scan on a 3T MRI system (Magnetom Allegra, Siemens, Erlangen, Germany) using an echo planar imaging (EPI) sequence with the following parameters: TR = 2080 ms, TE = 30 ms, 32 axial slices parallel to AC-PC line, matrix = 64 x 64, in plane resolution = 3x3 mm^2, slice thickness = 2.5 mm, 50% skip, flip angle = 70°. Resting scans lasted for 7 minutes and 20 seconds for a total of 220 volumes. A T1-weighted three-dimensional modified driven equilibrium Fourier transform scan (Deichmann et al., 2004) was acquired for each subject for anatomical localization purposes and for grey matter volumetry; the parameters were as follows: TR = 1338 ms, TE = 2.4 ms, TI = 910 ms, flip angle = 15°, matrix = 256 x 224 x 176, FOV = 256 x 224 mm^2, slice thickness = 1 mm, total scan time = 12 min.

2.3 DATA PREPROCESSING

Connectivity toolbox (CONN: functional connectivity toolbox, Whitfield-Gabrieli and Nieto-Castanon, 2012) was used to preprocess functional images. Images were slice-time corrected, realigned to the first image and coregistered to the high resolution T1 image using the mean EPI. Coregistered volumes were then normalized into Montreal Neurological Institute space coordinates (voxel size: 2x2x2 mm^3) and smoothed using an 8x8x8 mm^3 full width at half maximum Gaussian kernel. T1 image was also segmented to separate brain tissue components, i.e., gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). Before any statistical analysis, data underwent physiological noise reduction consisting in regressing out the six parameters of realignment as well as the first five PCA eigenvectors of the EPI timecourse over CSF and WM (Behzadi et al., 2007). Data were then detrended and filtered in the frequency range 0.01-0.073 Hz.

2.4 FUNCTIONAL CONNECTIVITY

FC maps were obtained via seed-to-voxel correlation analysis with seed in a-priori PCC. Briefly, a Pearson correlation coefficient was computed between mean-seed time-course and all other voxel time-courses; subsequently, a z-Fisher transformation was applied. The choice of the seed was motivated by the fact that AD-related changes in FC reported in literature were mainly found within DMN (Greicius et al., 2004), which is well known to have a key node in PCC (Greicius et al., 2003a).

2.5 ALFF

Fast Fourier Transform (FFT) algorithm (3dPeriodogram; AFNI package, Cox, 1996) was used to transform the filtered time series of each voxel in the frequency domain in order to obtain the corresponding power spectral density. The latter was square rooted and averaged across frequencies. ALFF maps of each subject were standardized into *z*-scores (Zuo et al., 2010).

2.6 GRAY MATTER VOLUMETRY

Modulated gray matter probability maps, obtained from the segmentation step, were used as a measure of voxel-wise gray matter volume (GMV).

2.7 ROI-BASED ANALYSIS

Cerebral cortex was parceled in 47 a-priori ROIs according to Brodmann classification (WFUPickAtlas MATLAB toolbox). Mean values of FC, ALFF and GMV were extracted for each ROIs. Group comparisons were carried out via two-sample, two-tailed *t*-tests.

3 RESULTS

Figure 1 shows the group results for FC and ALFF. At the whole-brain level, both measures slightly decrease in AD relative to HC (average *z* decreases 0.04±0.01 and 0.03±0.01 for FC and ALFF, respectively). However, FC decreases in the brain of AD patients, while the patterns of ALFF changes are apparently more complex. In particular, ALFF increases in frontal areas and decreases in parietal areas.

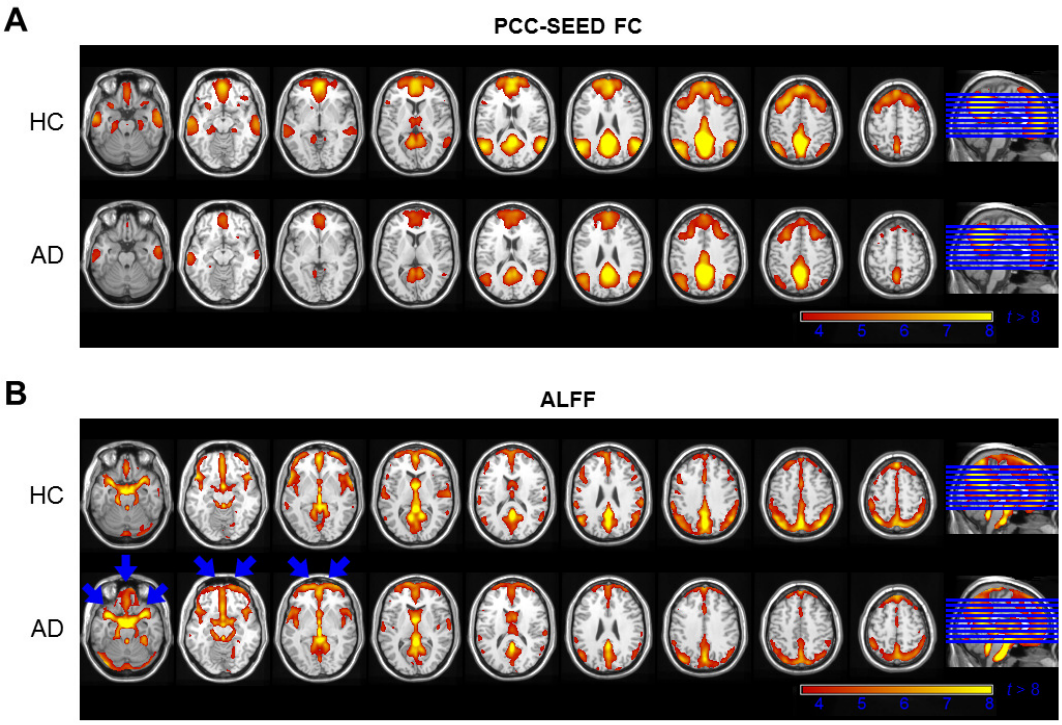


Figure 1. Group results for FC (A) and ALFF (B) obtained via one-sample *t*-tests ($p<0.001$, unc.). HC are shown on top and AD patients are shown on bottom of each panel. At the whole-brain level both FC and ALFF decrease in AD relative to HC. Nonetheless, FC broadly decreases while ALFF increases in frontal areas and decreases in parietal areas. The increase in ALFF is especially prominent in orbitofrontal portion of the prefrontal cortex and medial temporal lobe (blue arrows; see text).

Table 1. Brodmann areas exhibiting significant FC decreases in patients versus controls.

BA	Name	p [1]	ΔFC [2]	ΔGMV [3]	Function
11	ventromedial prefrontal cortex	0.002	-0.09±0.03	-22±4%	cognitive empathy, motivation
8	frontal cortex	0.004	-0.12±0.04	-24±5%	working-memory
10	anterior prefrontal cortex	0.008	-0.09±0.03	-27±4%	working-memory, attention
21	middle temporal gyrus	0.009	-0.10±0.03	-23±3%	language
38	temporal pole	0.029	-0.07±0.03	-24±4%	language, emotion, executive functions, memory
9	dorsolateral/medial prefrontal cortex	0.043	-0.06±0.03	-20±4%	planning, motivation, working-memory
35	perirhinal cortex	0.044	-0.09±0.05	-6±3%	semantic memory
31	dorsal posterior cingulate cortex	0.048	-0.04±0.02	-16±3%	working-memory, awareness
39	angular gyrus	0.049	-0.08±0.04	-23±4%	language, semantic memory

[1] Two-sample, two-tailed t-test, sorted for p value (up to a significance level of 0.05) for pathology-induced FC changes. Of note, no enhancement was found.
[2] Fisher z change and relative SEM between FC in HC and AD.
[3] Percent change and relative SEM between GMV in HC and AD.

Table 1 shows the areas displaying significant FC decreases in AD patients compared with healthy controls, along with GMV changes. These areas of the DMN are involved in cognitive functions and are known to be affected by the disease. In particular, PCC and PFC are characterized by high density of β-amyloid and neurofibrillary tangles (Chu et al., 1997). Among the areas characterized by connectivity loss are those known to be early affected by AD, including temporal pole (Ding et al., 2009) and perirhinal cortex (Didic et al., 2011). Remarkably, the decrease in FC is paralleled by a widespread reduction in GMV, consistent with the substantial cortical atrophy that characterize AD (Chetelat et al., 2010). To rule out any bias related to the choice of the seed, we performed a voxel-wise connectivity measure (AFNI package, Cox, 1996) with no a-priori assumption (Cole et al., 2010). Essentially the same areas are found when considering a global measure of functional connectivity instead of the seed-based measure (data not shown), suggesting that DMN connectivity governs cortical FC patterns.

Table 2 shows the AD-related changes in the amplitude of the fluctuations that are used to calculate FC. We found that ALFF increases in some regions and decreases in others, without a clear-cut relationship with FC (i.e., no pairwise correlation survived statistical threshold). Interestingly enough, increased ALFF occurs even in areas of significantly reduced GMV. Indeed, areas characterized by suppressed ALFF include regions early affected by AD, such as postcentral gyrus (Stephen et al., 2010), supramarginal gyrus (Harasty et al., 1999) and precuneus (Ross et al., 1996), besides the already cited posterior cingulate and perirhinal cortex. Similarly, ALFF enhancement occurs in areas characterized by high density of neurofibrillary tangles (Chu et al., 1997) including one of the very first affected region by AD, namely entorhinal cortex (Khan et al., 2014), which is distinguished by a severe neuronal loss (Gomez-Isla et al., 1996). According to our analysis, there are only two common areas to the significant changes in either FC or ALFF, namely the ventromedial prefrontal cortex (vmPFC; Brodmann area 11) and the dorsal posterior cingulate cortex (dPCC; Brodmann area 31). Both these areas exhibit a loss of connectivity within the DMN (Figure 2A). Yet, the slow oscillations underlying this decrease in FC are suppressed for dPCC but enhanced for vmPFC (Figure 2B,C).

4 DISCUSSION

The clinical course of AD reflects the progressive neuropathology leading to irreversible cognitive deficits, such as impairment of learning and memory as well as language difficulties. In agreement with previous results, we found that the pathology affects primarily cognitive areas and especially the DMN (Beason-Held, 2011). The cognitive decline accompanying AD appears to explain the decrease in ALFF that we found mainly in cortical areas involved in cognitive functions. In particular, ALFF suppression may reflect hypo-activity and cognitive impairment associated with cortical atrophy (Francis et al., 1993).

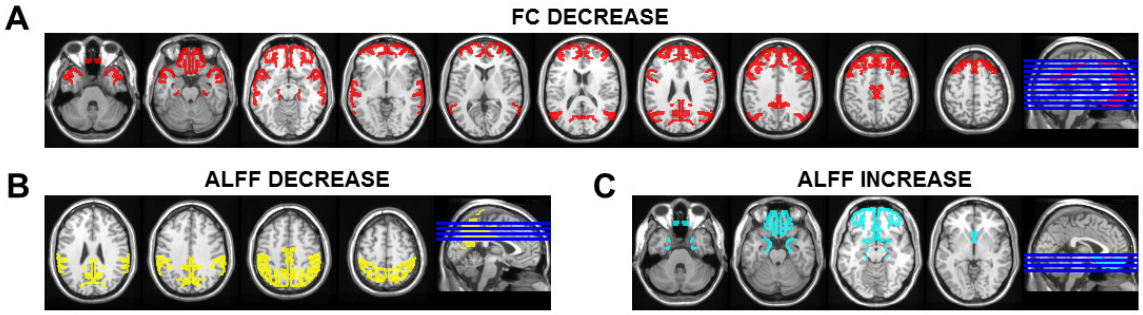


Figure 2. Summary of the main findings of ROI-based analysis underlying changes in AD patients compared with healthy controls. Brain regions (aggregated Brodmann's areas) exhibiting (A) decrease in functional connectivity and either (B) decrease or (C) increase in low-frequency fluctuations amplitude. Areas of FC decrease include the DMN hubs of PCC and PFC. Noticeably, ALFF changes in opposite directions in these two regions, decreasing in PCC and increasing in PFC. While no region exhibits increase in FC, alterations in ALFF separate frontal/temporal from parietal/occipital areas (the only exception being the perirhinal cortex, which is below the slices shown in panel B).

However, we also found cortical areas exhibiting increased ALFF. It is possible that the latter areas account for the restless activity commonly observed during dementia of the Alzheimer type (Devanand et al., 1997), although there is an apparent mismatch between the potentially involved regions (Banno et al., 2014).

An intriguing possibility to explain ALFF increments comes from the fact that deposition of β-amyloid induces neuronal hyperexcitability specifically in the frontal cortex (Kellner et al., 2014). Although the latter finding has been reported in animal models of AD, it is consistent with the observation that temporofrontal epileptic syndromes are more likely in AD patients than in general elderly population (e.g., Garcia-Cabrero et al., 2013). Alternatively, it is possible that the higher ALFF in prefrontal regions of AD patients compared with healthy subjects might reflect a deeper state of alertness/anxiety experienced by these patients while in the MR scanner.

Table 2. Brodmann areas exhibiting significant ALFF changes in patients versus controls.

BA	Name	p [1]	ΔALFF [2]	ΔGMV [3]	Function
Suppressed					
31	dorsal posterior cingulate cortex	0.001	-0.33±0.09	-16±3%	working-memory, awareness
40	supramarginal gyrus	0.001	-0.27±0.08	-25±3%	language processing
36	perirhinal cortex	0.004	-0.30±0.10	-20±4%	context-free (semantic) memory
7	precuneus/superior parietal lobule	0.005	-0.24±0.08	-29±3%	episodic memory and language
2	postcentral gyrus	0.028	-0.23±0.10	-20±3%	somatosensory processing
23	ventral posterior cingulate cortex	0.041	-0.28±0.13	-3±5%	working-memory, awareness
Enhanced					
11	ventromedial prefrontal cortex	0.007	+0.28±0.09	-22±4%	motivation, cognitive empathy
34	dorsal entorhinal cortex	0.022	+0.30±0.13	-32±4%	spatial memory
28	ventromedial entorhinal cortex	0.043	+0.23±0.11	-32±3%	
25	subgenual cortex	0.049	+0.38±0.20	-11±5%	mood, anxiety

[1] Two-sample, two-tailed *t*-test, sorted for *p* value (up to a significance level of 0.05) for pathology-induced ALFF changes.
[2] Z-score change and relative SEM between ALFF in HC and AD.
[3] Percent change and relative SEM between GMV in HC and AD.

Overall, our results indicate that in principle hypo- and hyper-activity can coexist in the brain of AD patients. Nonetheless, the fact that substantial cortical atrophy, a marker of neuronal loss, coexists with increased oscillations amplitude suggest that physiological mechanisms besides neuronal activity may contribute to the generation of these slow oscillations. Potentially causal mechanisms involve fluctuating hemodynamic variables, such as arterial blood pressure, intracranial pressure, microcirculation and cerebral oxygenation (de la Torre, 2010). Remarkably, astrocytes plays an important role in the regulation of cerebral perfusion. We have recently proposed that astrocytes could represent a possible origin for the slow BOLD oscillations, a process mediated by Ca^{2+}-dependent signaling and in particular by astrocytic Ca^{2+} waves (DiNuzzo et al., 2011). Indeed, astrocytic hyperactivity (e.g., reactive astrocytosis) and subsequent alteration of Ca^{2+} waves oscillations have been implicated in the pathogenesis of AD (see Chow et al., 2010 and references therein), likely in a region-dependent manner (Grolla et al., 2013).

The pattern of ALFF changes that we report here are in close qualitative agreement with previous work (Wang et al., 2011). Specifically, we found that ALFF suppression mainly involves a set of parietal (middle-central and lateral) and posterior temporal areas including posterior cingulate cortex, precuneus, as well as postcentral and supramarginal gyri. We also found that ALFF enhancement is circumscribed to a set of frontal and anterior temporal areas including ventromedial prefrontal cortex, as well as entorhinal and subgenual cortices. One single exception is represented by the perirhinal cortex, a region bounded by entorhinal cortex, which contrary to the latter exhibits a decrease in ALFF. This findings is somewhat surprising considering the anatomical contiguity and the early involvement of both areas in the pathology (Juottonen et al., 1998). However, entorhinal and perirhinal cortices are functionally distinct and there is very low reciprocal activity propagation between them, likely because of their role in gating information between neocortex and hippocampus (de Curtis and Pare, 2004).

5 CONCLUSION

In the present work, we examined the changes in FC and ALFF occurring in specific brain areas of AD patients relative to healthy subjects. Using a ROI-based approach to increase sensitivity, we confirmed previous works showing that FC and GMV are both reduced in AD. We found some brain areas exhibiting increased ALFF on top of loss of connectivity and cortical atrophy. Notably, posterior cingulate cortex and medial prefrontal cortex, two central hubs of the DMN, are characterized by decreased FC yet opposite changes in ALFF. Our results arguably indicate that FC and ALFF may vary independently on each other. As such, the relation between synchronization and amplitude of low-frequency fluctuations might be complex. In fact, a loss of signal correlation could have been explained as the result of decreased signal amplitude towards the noise level. At odds with this simple explanation, we found that reduced FC can occur even in areas with increased ALFF. This is especially relevant in vmPFC, which is characterized by the most significant FC loss and ALFF gain. Further research is required to establish the physiological mechanisms underlying the regional difference in ALFF, and whether it does preexist to, or is induced by, the pathology.

REFERENCES

Banno K, Nakaaki S, Sato J, Torii K, Narumoto J, Miyata J, Hirono N, Furukawa TA, Mimura M, Akechi T (2014) Neural basis of three dimensions of agitated behaviors in patients with Alzheimer disease. Neuropsychiatric disease and treatment 10:339-348.

Beason-Held LL (2011) Dementia and the default mode. Current Alzheimer research 8:361-365.

Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37:90-101.

Braak H, Braak E, Bohl J (1993) Staging of Alzheimer-related cortical destruction. European neurology 33:403-408.

Chetelat G, Villemagne VL, Bourgeat P, Pike KE, Jones G, Ames D, Ellis KA, Szoeke C, Martins RN, O'Keefe GJ, Salvado O, Masters CL, Rowe CC (2010) Relationship between atrophy and beta-amyloid deposition in Alzheimer disease. Ann Neurol 67:317-324.

Chow SK, Yu D, Macdonald CL, Buibas M, Silva GA (2010) Amyloid beta-peptide directly induces spontaneous calcium transients, delayed intercellular calcium waves and gliosis in rat cortical astrocytes. ASN Neuro 2:e00026.

Chu CC, Tranel D, Damasio AR, Van Hoesen GW (1997) The autonomic-related cortex: pathology in Alzheimer's disease. Cereb Cortex 7:86-95.

Claus JJ, Ongerboer de Visser BW, Walstra GJ, Hijdra A, Verbeeten B, Jr., van Gool WA (1998) Quantitative spectral electroencephalography in predicting survival in patients with early Alzheimer disease. Arch Neurol 55:1105-1111.

Cole MW, Pathak S, Schneider W (2010) Identifying the brain's most globally connected regions. Neuroimage 49:3132-3148.

Cook IA, Leuchter AF (1996) Synaptic dysfunction in Alzheimer's disease: clinical assessment using quantitative EEG. Behavioural brain research 78:15-23.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162-173.

de Curtis M, Pare D (2004) The rhinal cortices: a wall of inhibition between the neocortex and the hippocampus. Prog Neurobiol 74:101-110.

de la Torre JC (2010) The vascular hypothesis of Alzheimer's disease: bench to bedside and beyond. Neuro-degenerative diseases 7:116-121.

Deichmann R, Schwarzbauer C, Turner R (2004) Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. Neuroimage 21:757-767.

Devanand DP, Jacobs DM, Tang MX, Del Castillo-Castaneda C, Sano M, Marder K, Bell K, Bylsma FW, Brandt J, Albert M, Stern Y (1997) The course of psychopathologic features in mild to moderate Alzheimer disease. Archives of general psychiatry 54:257-263.

Didic M, Barbeau EJ, Felician O, Tramoni E, Guedj E, Poncet M, Ceccaldi M (2011) Which memory system is impaired first in Alzheimer's disease? Journal of Alzheimer's disease : JAD 27:11-22.

Ding SL, Van Hoesen GW, Cassell MD, Poremba A (2009) Parcellation of human temporal polar cortex: a combined analysis of multiple cytoarchitectonic, chemoarchitectonic, and pathological markers. The Journal of comparative neurology 514:595-623.

DiNuzzo M, Gili T, Maraviglia B, Giove F (2011) Modeling the contribution of neuron-astrocyte cross talk to slow blood oxygenation level-dependent signal oscillations. J Neurophysiol 106:3010-3018.

Francis PT, Sims NR, Procter AW, Bowen DM (1993) Cortical pyramidal neurone loss may cause glutamatergic hypoactivity and cognitive impairment in Alzheimer's disease: investigative and therapeutic perspectives. J Neurochem 60:1589-1604.

Garcia-Cabrero AM, Guerrero-Lopez R, Giraldez BG, Llorens-Martin M, Avila J, Serratosa JM, Sanchez MP (2013) Hyperexcitability and epileptic seizures in a model of frontotemporal dementia. Neurobiol Dis 58:200-208.

Gili T, Cercignani M, Serra L, Perri R, Giove F, Maraviglia B, Caltagirone C, Bozzali M (2011) Regional brain atrophy and functional disconnection across Alzheimer's disease evolution. J Neurol Neurosurg Psychiatry 82:58-66.

Gomez-Isla T, Price JL, McKeel DW, Jr., Morris JC, Growdon JH, Hyman BT (1996) Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. J Neurosci 16:4491-4500.

Greicius MD, Krasnow B, Reiss AL, Menon V (2003a) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci U S A 100:253-258.

Greicius MD, Krasnow B, Reiss AL, Menon V (2003b) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci U S A 100:253-258.

Greicius MD, Srivastava G, Reiss AL, Menon V (2004) Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci U S A 101:4637-4642.

Grolla AA, Sim JA, Lim D, Rodriguez JJ, Genazzani AA, Verkhratsky A (2013) Amyloid-beta and Alzheimer's disease type pathology differentially affects the calcium signalling toolkit in astrocytes from different brain regions. Cell death & disease 4:e623.

Grunwald M, Busse F, Hensel A, Riedel-Heller S, Kruggel F, Arendt T, Wolf H, Gertz HJ (2002) Theta-power differences in patients with mild cognitive impairment under rest condition and during haptic tasks. Alzheimer disease and associated disorders 16:40-48.

Harasty JA, Halliday GM, Kril JJ, Code C (1999) Specific temporoparietal gyral atrophy reflects the pattern of language dissolution in Alzheimer's disease. Brain 122 (Pt 4):675-686.

He Y, Wang L, Zang Y, Tian L, Zhang X, Li K, Jiang T (2007) Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study. Neuroimage 35:488-500.

Jeong J (2004) EEG dynamics in patients with Alzheimer's disease. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 115:1490-1505.

Juottonen K, Laakso MP, Insausti R, Lehtovirta M, Pitkanen A, Partanen K, Soininen H (1998) Volumes of the entorhinal and perirhinal cortices in Alzheimer's disease. Neurobiology of aging 19:15-22.

Kellner V, Menkes-Caspi N, Beker S, Stern EA (2014) Amyloid-beta alters ongoing neuronal activity and excitability in the frontal cortex. Neurobiology of aging 35:1982-1991.

Khan UA, Liu L, Provenzano FA, Berman DE, Profaci CP, Sloan R, Mayeux R, Duff KE, Small SA (2014) Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. Nat Neurosci 17:304-311.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34:939-944.

Ross SJ, Graham N, Stuart-Green L, Prins M, Xuereb J, Patterson K, Hodges JR (1996) Progressive biparietal atrophy: an atypical presentation of Alzheimer's disease. J Neurol Neurosurg Psychiatry 61:388-395.

Selkoe DJ (1994) Cell biology of the amyloid beta-protein precursor and the mechanism of Alzheimer's disease. Annual review of cell biology 10:373-403.

Stephen JM, Montano R, Donahue CH, Adair JC, Knoefel J, Qualls C, Hart B, Ranken D, Aine CJ (2010) Somatosensory responses in normal aging, mild cognitive impairment, and Alzheimer's disease. Journal of neural transmission (Vienna, Austria : 1996) 117:217-225.

Takashima A (2009) Amyloid-beta, tau, and dementia. Journal of Alzheimer's disease : JAD 17:729-736.

Wang Z, Yan C, Zhao C, Qi Z, Zhou W, Lu J, He Y, Li K (2011) Spatial patterns of intrinsic brain activity in mild cognitive impairment and Alzheimer's disease: a resting-state functional MRI study. Hum Brain Mapp 32:1720-1740.

Whitfield-Gabrieli S, Nieto-Castanon A (2012) Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect 2:125-141.

Zhang HY, Wang SJ, Liu B, Ma ZL, Yang M, Zhang ZJ, Teng GJ (2010) Resting brain connectivity: changes during the progress of Alzheimer disease. Radiology 256:598-606.

Zuo X-N, Martino AD, Kelly C, Shehzad ZE, Gee DG, Klein DF, Castellanos FX, Biswal BB, Milham MP (2010) The oscillating brain: complex and reliable. Neuroimage 49:1432-1445.

FRACTAL ANALYSIS OF BREAST RADIOLOGICAL IMAGES

Alessandro Palombo[1*], Maria Laura Luciani[2], Federica Pediconi[2], Mariano Bizzarri[1]

[1]*Dept. of Experimental Medicine, "Sapienza" University of Rome, Systems Biology Group Lab, Via A. Scarpa 14, 00161 Rome, Italy. *Corresponding author. alessandro.palombo@uniroma1.it; mariano.bizzarri@uniroma1.it.* [2] *Dept. of Radiological sciences, Oncology and Anatomical Pathology, "Sapienza" University of Rome, Viale Regina Elena 324, 00161 Rome, Italy. marialauraluciani@gmail.com; federica.pediconi@uniroma1.it.*

Keywords: Fractal analysis; breast cancer; radiology.

Abstract Masses due to benign breast diseases and tumors due to breast cancer present significantly different shapes on mammograms. In general, malignant tumors appear with rough, spiculated and complex boundaries or contours, whereas benign masses present smooth, round, or oval contours. Additionally breast cancer tissues display architectural distortion and abnormalities in tissue texture. Fractal analysis may be used to derive shape features to perform pattern classification of breast masses and tumors, thus improving diagnostic accuracy.

1 INTRODUCTION

Interest in fractal geometry has been increasing steadily during the last decades, given that such approach may extract more information embedded into morphological data and, furthermore, can link them to the non-linear dynamics which lies behind the behavior of complex systems (Shelhamer, 2007). In turn, non-linear dynamics, according to non-equilibrium thermodynamics, may enact chaotic behaviors, constraining natural phenomena into fractal-like structures. Physiologists and physicians have only recently begun to quantify the widespread occurrence of such possibilities in Biology, thus casting on doubt the long-held principle of 'homeostasis' on which, according to the Cannon's seminal work, the edifice of Physiology has been build.

Mandelbrot (Mandelbrot, 1975) introduced the term "fractal" (from the Latin *fractus*, meaning "broken") to characterize spatial or temporal phenomena that are continuous, but not differentiable. Fractal dimension differs from our intuitive notion of dimension in that it can be a non-integer value, and the more irregular and complex an object is, the higher its fractal dimension. Fractals are in some

way the product of non-linear dynamical processes: a process characterized by non-linear dynamics and behaving far from the equilibrium will likely shape its environment (a cloud, a biological rhythm, a tissue profile, an enzymatic reaction) by leaving behind a fractal. A fractal is a geometrical set that exhibits a repeating pattern that displays at every scale, a property called self-similarity. Fractal geometry has the ability to quantify the irregularity and complexity of objects with a measurable value called the fractal dimension. Fractal dimension (FD) differs from our intuitive notion of dimension in that it can be a non-integer value and the more irregular and complex an object is, the higher its fractal dimension. The FD provides a quantitative measure about how an irregular geometric body fills its space. In this respect, according to the Euclidean geometry, lines have a FD of one, squares have two dimensions, whereas cubes have three. In all these cases – a line, a square, a cube – the geometrical figure entirely fills its dimensional space. However fractals have non integer (fractional) dimensions and therefore they do not fill precisely their space: while a smooth Euclidean line has FD=1, a fractal line has a FD between one and two. The greater the FD, the greater the chance that the fractal will fill the entire space. FD, according to Mandelbrot, is defined by a set of values of c in the complex plane for which the orbit of 0 under iteration of the complex quadratic polynomial remains $Z_{n+1} = Z_n^2 + C$ bounded. That is, a complex number c is part of the Mandelbrot set if, when starting with $z_0 = 0$ and applying the iteration repeatedly, the absolute value of zn remains bounded however large n gets. For example, letting c = 1 gives the sequence 0, 1, 2, 5, 26,…, which tends to infinity. As this sequence is unbounded, 1 is not an element of the Mandelbrot set. On the other hand, c = −1 gives the sequence 0, −1, 0, −1, 0,…, which is bounded, and so −1 belongs to the Mandelbrot set.

Actually, the self-similarity (fractal) dimension FD is defined as follows (Peitgen et al., 2004).Consider a self-similar pattern that exhibits a number of self-similar pieces N. This figure is of actual length L, which can be decomposed into N segments of length S, such that L = N x S. The minimum distance (the linear distance Lx) between the ends of the segment will instead be given by the minimum number of segments M (magnification factor) that satisfies Lx = M x S. the ratio L / Lx provides the fractal value, expressing how the actual value differs from the linear one. In fact, if N> M, we will have $N = M^D$: 'D' is indeed the fractal value and can be obtained by plotting the log ($N=M^D$) vs. Log M, thus obtaining:

$$\log (M^D)/\log M \text{ and } D = \log (M)/\log (M). \qquad (1)$$

Alternatively we can consider that the distance L covered by a self-similar pattern that exhibits a number N of self-similar pieces of length S, has a reduction factor equal to L/S. Therefore the power law to be satisfied is $N=1/S^D$, where D=1 for a linear distance. It follows that D= log (N)/log (1/S), known as the Hausdorff-Besicovitch definition of fractal dimension (Iftekharuddin et al., 2003). The D value indicates how many times the N segment must be repeated in order to cover the distance L. The greater the D value, the greater N will be. As N grows, the space will be more filled by the picture (Fig.1).

Additionally, the fractal dimension has a thermodynamic meaning and can be viewed as an intensive measure of the "overall" (morphologic) complexity (Smith et al., 1996; Cutting and Garvin 1987). Like many summary statistics (e.g., mean), the fractal dimension is obtained by .averaging. variation in data structure. In doing so, information is necessarily lost (Normant and Tricot 1991). The estimated fractal dimension of a lakeshore, for example, tells us nothing about the actual size or overall shape of the lake, nor can we reproduce a map of the lake from the fractal dimension alone. However, the fractal dimension does tell us a great deal about the relative complexity of the lakeshore and as such is an important descriptor when used in conjunction with other measures. It should be emphasized that the fractal dimension is a descriptive, quantitative measure; it is a statistic, representing an attempt to estimate a single-valued number for a property (complexity) of an object

with a sample of data from the object. One can view the fractal dimension in much the same way that thermodynamics might view intensive measures as temperature. In other words, fractal dimension can be considered a systems property and, together with one or more independent variables, could enable constructing a diagram of phases, like that relying on temperature, pressure, and volume for gas/liquid/solid phase transitions.

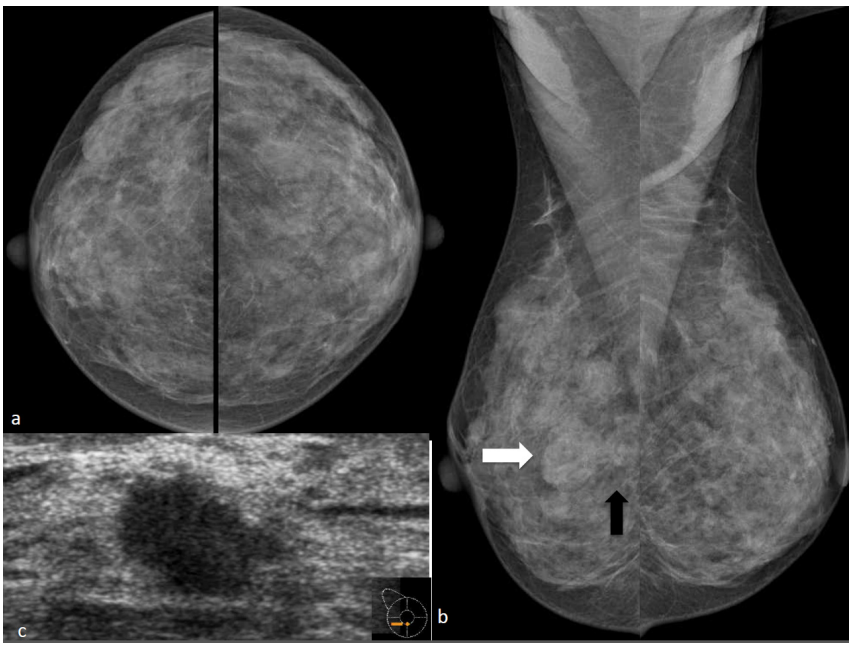


Figure 1: 45 yo patient routine screening a)Direct digital XR mammography: bilateral CC (cranio-caudal) and MLO (medio-lateral oblique) projections revealed dense breast. b)RMLO (right medio-lateral oblique) projection: white arrow, round and well defined margins opacity (benign finding) at the same site ultrasound detected a benign cyst; black arrow, architectural distortion (malignant finding) c)Ultrasound: hypoechoic nodule (malignant finding) between lowers quadrants of the right breast (same site of the architectural distortion); at pathology it was proved to be IDC (invasive ductal carcinoma).

Many physiological and pathological processes display rhythms, dynamics or architecture morphology characterized by a fractal structure (Losa et al., 2002; Cross 1997). In the human body fractal-like structure abound in networks of blood vessels, nerve, ducts and bronchial tree configuration. Moreover, cells, cytoskeleton microtubule configuration, as well as the branching pattern of several tissues (as the His-Purkinje system), are characterized by a fractal architecture characterizes. Although these fractal anatomies serve apparently disparate functions in different tissue and organ systems, several common anatomical and physiological themes emerge. Fractal branches or folds greatly amplify the surface area available for absorption (as in intestine and in the alveolar tree), distribution or collection (by the blood vessels, bile ducts and bronchial tubes), and information processing (by the nerves) (Landini 2002). Fractal structures are robust and more resistant to injury, and allow tissue to maintain their connectedness and function despite extensive damage. In addition, higher fractal values of the outer leaflet of the cell membrane is associated to increased motility and invasiveness, a property that embryonic cells share with cancer cells (Pasqualato et al., 2013).

Indeed, compelling evidence demonstrate that cancerous cells and tissues display higher fractal values (Baish and Jain 2000). Fractal analysis helped in discriminating benign from malignant neoplasms (Cross et al., 1995), low from high grade tumours (Claridge et al., 1992), and invasive

from non-invasive cancers (Pasqualato et al., 2012). Furthermore, fractal studies elucidated some aspects of the complex interplay between cancer cells and stroma by suggesting that tumour vascular architecture is determined by heterogeneity in the cellular interaction with the extracellular matrix rather than by simple gradients of diffusible angiogenic factors (Gazit et al., 1995). Moreover, fractal analysis of the interface between cancer and normal tissues helps in understanding how cell detachment from the primary mass and infiltration into adjacent tissue occurs through a non-mutational mechanism (Michaelson et al., 2005). These statements are of outmost relevance by keeping in mind that, for now, despite the amazing growth in our understanding of the molecular mechanisms of cancer, diagnosis is mostly still done by visual examination of images and by the morphological examination of radiological pictures, microscopy of cell and tissues, and so forth (Rosai, 2001). Thereby, a quantitative and operationally reproducible approach, such that provided by fractal analysis, could lead to a remarkable improvement in both cyto-histological and radiographic diagnostic accuracy.

2 FRACTALS IN BREAST RADIOLOGY

Despite the progress made during the last decades, radiological examinations still suffer from a lack of accuracy that limits its clinical reliability, thus requiring a set of different diagnostic tools in order to achieve a proper assessment needed in setting an appropriate pretreatment planning. For instance, the sensitivity of mammography in detecting breast cancer ranges from 63% to 98% (Burhenne et al., 1994; Kerlikowske et al., 1996) and has been reported to be as low as 30%–48% in dense breasts (Margaret et al., 2000; Kolb et al 2002). Furthermore, 27%–34% of breasts had additional malignant foci not seen mammographically (Moon et al., 2002; Hlawatsch et al., 2002; Luciani et al., 2011).

Namely, mammographic examination may both leads to false and positive results. False-positive results (i.e., erroneous evaluation of a breast mass as a cancer) are likely to occur in younger women, in women who have had previous breast biopsies, women with a family history of breast cancer, and women who are taking estrogen. False negative results (normal mammograms even though breast cancer is present) are mainly due to high breast density (Pediconi et al., 2009)

As connective components and/or fibro-glandular tissue displays density value close to the values of cancerous tissues, it becomes hard to distinguish the former form the latter, and in women with denser breasts proper clinical assessment is therefore biased. FIG1 False-negative results occur more often among younger women than among older women because younger women are more likely to have dense breasts. As a woman ages, her breasts usually become more fatty, and false-negative results become less likely. Hence, with the aim of improving the accuracy and efficiency of screening programs for the detection of early signs of breast cancer, a number of research projects are focusing on developing

methods for computer-aided diagnosis to assist radiologists in diagnosing breast cancer by image enhancement and image analysis (Peitgen, 2002). Many of those programs have been focused on quantifying the abnormalities recognized in radiological examination (mammography) of the breast.

Mammography allow to detect tumors by evidencing a) the presence of a dense mass, featured by irregular contours FIG2; b) change in the texture or distortion in the mammogram, usually reinforced by the presence of micro-calcification. Benign masses are usually homogeneous, whereas cancers display a heterogeneous texture. FIG3 In order to discriminate among benign and malignant masses a lot of different measures of texture and edge sharpness have been proposed (Mudigonda et al., 2001; Mudigonda et al. 2000), in some cases by suggesting a mix of morphological and texture parameters to classify breast tumors (Sahiner et al., 2001; Alto et al., 2005).

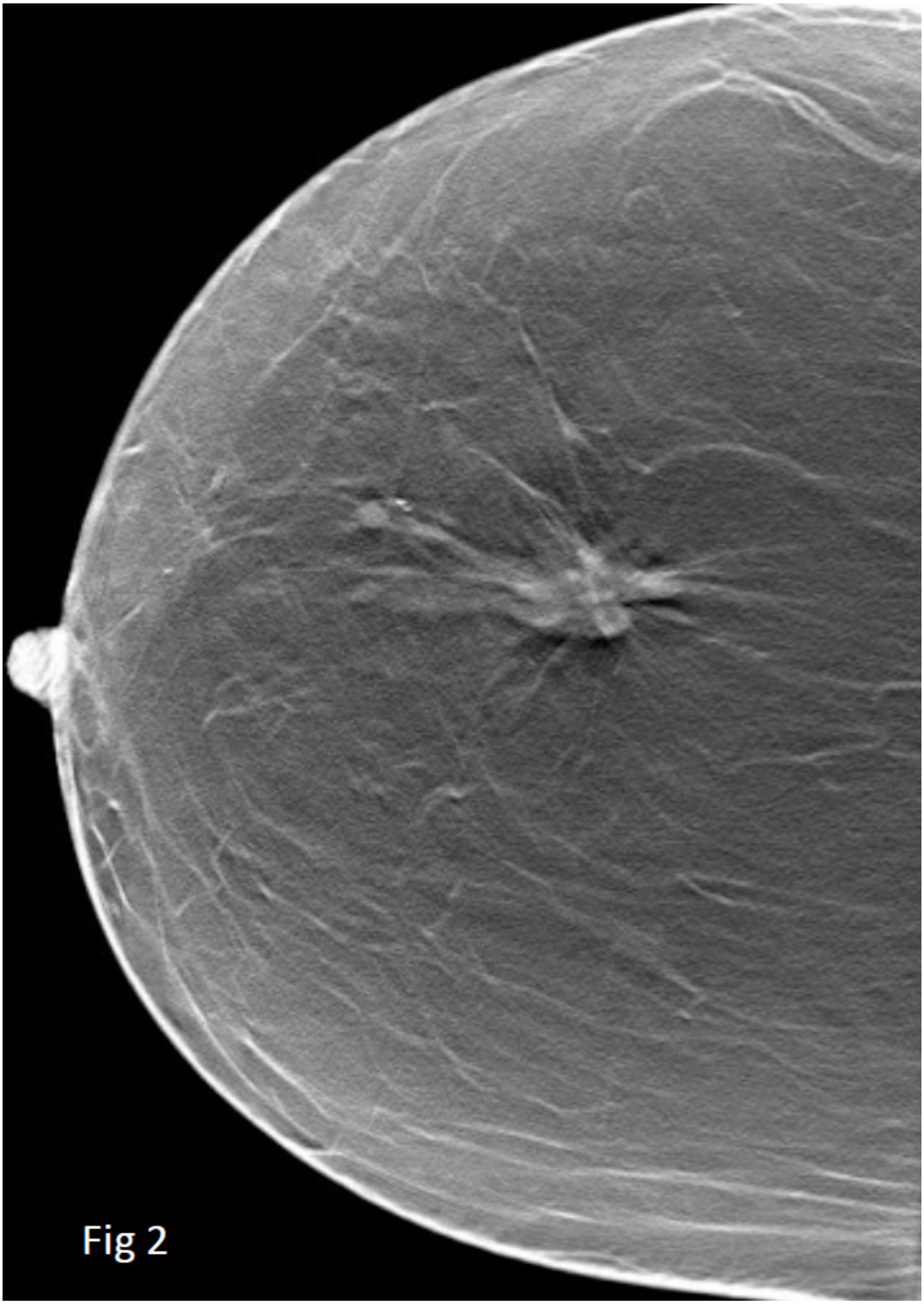


Figure 2: 3D tomosynthesis slice, right breast. Spiculated opacity (malignant finding).

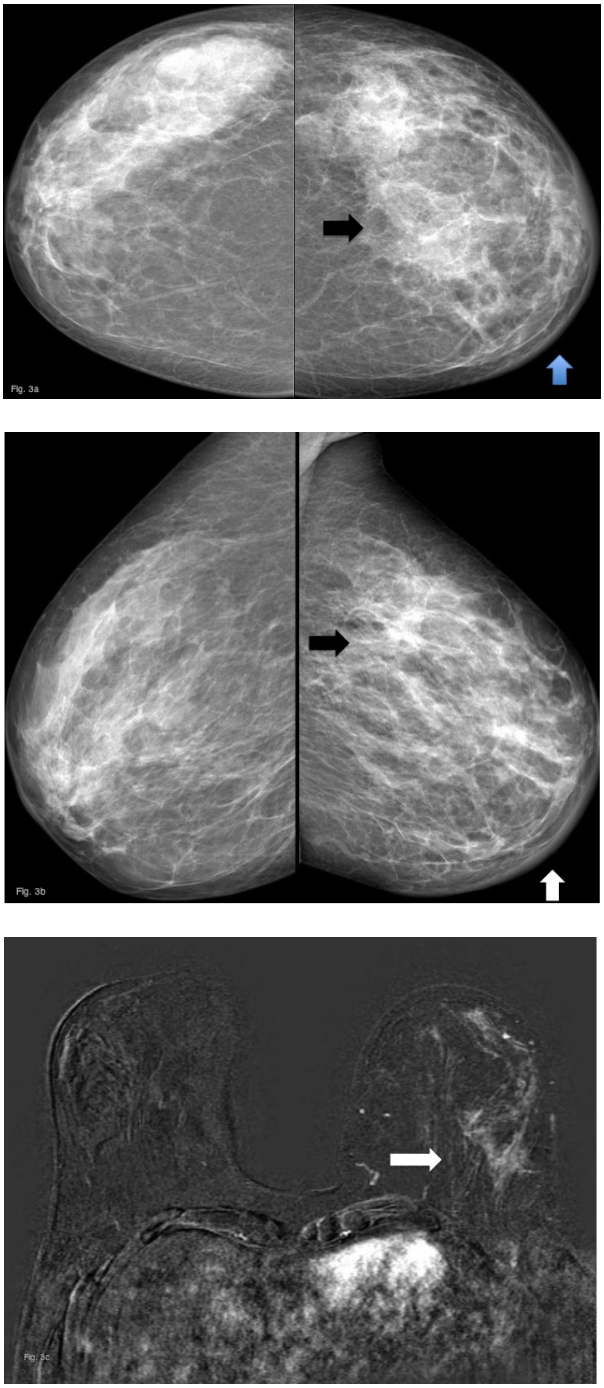


Figure 3: 50 yo patient without history of previous breast pathology or surgery. a-b) cranio-caudal and medio-lateral oblique projections: black arrow shows asymmetrical mammographic density on the left breast; blue and white arrows show skin thickening. They are both suspicious findings. c)3 T Breast MRI: VIBRANT contrast enhanced T1 weighted image confirms the suspicion of malignancy demonstrating a diffuse non mass like enhancement on the left breast.

Shape differences between benign and malignant tumors are noteworthy. Consequently, several form features such as compactness, fractional concavity, speculation index, Fourier-descriptor based factor (Rangayyan et al., 1997; Rangayyan et al., 2000) as well as fractal dimension (FD) have been proposed for their classification.

Fractal analysis has been successfully used in discriminating normal from malignant cancer cells on cytological studies (Bizzarri et al., 2011; Dey and Mohanty 2000; Ohri et al., 2004), evidencing also that malignant progression and invasiveness are characterized by a progressive increase in cell fractality (Pasqualato et al.,2012b), whereas the reversion of tumor phenotype is followed by an impressive change in both the form and the fractal dimension of the neoplastic cells (D'Anselmi et al., 2010). Unexpectedly, only a few investigations have been carried out to classify breast masses on the basis of the irregularity exhibited in their contours, meanwhile others scholars adopted a fractal approach to quantify the irregularity displayed by breast texture.

3 FRACTAL BREAST MASS SHAPE

Indeed, a typical benign mass has a contour that is round, smooth, and well defined, whereas a typical malignant tumor has a contour that is spiculated, rough, and ill-defined. These differences reflect different biological behavior and properties: a jagged contour is a prerequisite supporting invasiveness and migration, two main cancer features. Additionally, a speculate profile betray the higher complexity of cancer cell architecture, as FD may also be computed through entropy-based algorithms (Sankar and Thomas, 2009). Consequently, the significant differences between the boundary shape characteristics of benign masses and malignant tumors may be used to differentiate between them by deriving shape factors. A less cell-dense tissue usually has a lower fractal dimension than a tissue with a higher cell density (i.e., the number of cells divided by the tissue volume). When allowed to grow, the density of a tissue with a lower fractal dimension drops quickly. However, a tumor, since it has a higher fractal mass dimension, maintains a high density as it grows bigger, resulting in a more rapid growth rate and a larger final size. Fractal dimensions of infiltrating ductal adenocarcinomas of the breast are high (i.e., 2.98), which results in a very dense tissue compared with normal breast tissue (with a fractal dimension of about 2.25). As expected, the higher fractal dimension results in a high rate of growth (Norton, 2005).

Calculating the FD allowed Matsubara et al. to achieve a 96.5% accuracy (with 3.5% false positive results) in discriminating benign from malignant tumors of the breast. After a careful process of progressive exclusion (by using size, circularity, standard deviation and contrast analysis), "true" masses were identified and thus recognized as benign or malignant by means of fractal analysis of their profile (Matsubara et al., 1997). Quite similar results have been provided by Pohlman et al. (Pohlman et al., 1996), by V. Velanovich (Velanovich, 1998) (even if this Author reported a high value of false positive rate), as well as by Rangayyan and Nguyen (Rangayyan and Nguyen, 2007). This latter article demonstrated that FD of the breast mass may be obtained either directly from the 2-dimensional (2D) contour or from a 1-dimensional (1D) signature, by using different calculating methods (among which the box-counting and the ruler method). Fractal data were associated with four previously developed shape factors (speculation index, compactness, fractional concavity, Fourier Factor), The combination of fractal dimension with fractional concavity yields the highest trustworthiness (around 93%).

Fractal dimensions and fractal signatures are used to classify mammograms based on K means algorithm (Don and Revathy, 2007), meanwhile a few studies have suggested a conventional fractal modeling of mammograms to enhance the diagnostic evaluation of microcalcifications (Li, 1997). Since clusters of microcalcifications are one of the most important and often the only indicator for

malignant tumors, the reconstruction of the three-dimensional structure of these clusters from the two mammographic views, cranial-caudal and mediolateral-oblique, is critically important.

A very promising approach, based on the 2D Wavelet-Transform Modulus Maxima method was used to detect microcalcifications in mammograms and to characterize the fractal geometry of benign and malignant clusters; the Bayesian statistical analysis shows that, with 95% credibility, the probability that fractal breast lesions are malignant is between 74% and 98%. Alternatively, with 95% credibility, the probability that Euclidean breast lesions are benign is between 76% and 96% (Batchelder et al., 2014). These results support the notion that the fractal structure of malignant tumors is more likely to be associated with an invasive behavior into the surrounding tissue compared to the less invasive, Euclidean structure of benign tumors.

Microcalcifications study can be enhanced by taking the difference between the original image and the processed image. Bocchi et al. (Bocchi et al., 2004) used a fractal model to describe the mammographic background images and a matched filter to enhance microcalcifications against the background that is likely to be altered as a consequence of local, physical distortion induced by calcium aggregates. Hence, Tourassi et al. (Tourassi et al., 2006) applied fractal analysis to the investigation of architectural distortion of the regional microenvironment induced by calcification: FD, which was computed using the Fourier power spectrum method, enabled in discriminating the regions of interest depicting architectural distortion from those recognizing normal breast parenchyma. Rangayyan et al. (Rangayyan et al., 2008) reported methods for the detection of sites of architectural distortion in mammograms, by using phase portrait modeling, fractal dimension and texture features. This is a very important point, given that indirect signs of malignancy (such as architectural distortion, bilateral asymmetry, single dilated duct, and developing densities) are prominent features in almost 20% of breast cancers (Sickles, 2003). Moreover, as reported on a study dealing with false negative mammograms, improvement in the detection of architectural distortion could lead to significant improvement in the prognosis of breast cancer patients (Broeders et al., 2003).

4 CONCLUSIONS AND FUTURE PERSPECTIVES

Overall, data collected up to now confirm that fractal analysis is a useful tool in the classification of mammographic masses. The fractal dimension of benign masses is, in general, less than the fractal dimension of malignant tumors. Fractal analysis offers the opportunity to associate a quantitative assessment to the subjective, even skilled, diagnostic recognition performed by expert pathologists. The quantitative morphological appreciation leads to increased accuracy and diagnostic reliability, a result of priceless value in clinical setting. Yet, some methodological questions are still remains unanswered to date. At a first glance, the ruler method seems to provide a slightly more accurate classification when the 2D contour representation is considered. On the contrary, the 'traditional' box-counting method shows higher reliability in evaluating the 1D signature representation. However, the combined dataset confirms that the 1D signature representation consistently yields more accurate classification than the 2D contour representation (Nguyen and Rangayyan, 2005).

A widely recognized methodological drawback of that approach is that the mass profile was drawn by hand. Although this step is usually performed by an expert radiologist specialized in mammography, a potentially confounding bias cannot be excluded but only minimized. Even if a few studies (Sahiner et al., 2001) have shown that shape factors can perform well with automatically extracted contours, the subjective, hand-based intervention of the radiologist cannot be avoided.

It is tempting to speculate those limitations would benefit from the development of new modelling approaches and the improvement of semi-automatized methods.

REFERENCES

Alto, H., Rangayyan, R.M., Desautels, JEL. (2005) Content-based retrieval and analysis of mammographic masses. J Electron Imaging 14(2): 1-17.

Baish, J.W., Jain, R.K. (2000) Fractals and cancer. Cancer Res, 60: 3683–3688.

Batchelder, K.A., Tanenbaum, A.B., Albert, S. (2014). Wavelet-Based 3D Reconstruction of Microcalcification Clusters from Two Mammographic Views: New Evidence That Fractal Tumors Are Malignant and Euclidean Tumors Are Benign. Plos ONE, 9: e107580.

Bizzarri, M., Giuliani, A., Cucina, A., D'Anselmi, F., Soto, A.M., Sonnenschein, C., 2011. Fractal analysis in a systems biology approach to cancer. Semin. Cancer Biol. 21, 175e182.

Bocchi, L., Coppini, G., Nori, J., Valli, G. (2004) Detection of singleand clustered microcalcifications in mammograms using fractal models and neural networks. Med Eng Phys 26:303–312.

Broeders, M.J.M., Onland-Moret, N.C., Rijken, H.J.T.M., Hendriks, J.H.C.L, Verbeek, A.L.M., Holland, R. (2003) Use of previous screening mammograms to identify features indicating cases that would have a possible gain in prognosis following earlier detection. Eur J Cancer 39:1770–1775.

Burhenne HJ, Burhenne LW, Goldberg F, et al. (1994) Interval breast cancers in the Screening Mammography Program of British Columbia: analysis and classification. AJR Am J Roentgenol 162:1067–1071.

Claridge, E., Hall, P.N., Keefe, M., et al. (1992) Shape analysis for classification of malignant melanoma. J. Biomed. Eng. 14: 229–324.

Cross, S.S. (1997) Fractals in pathology. J Pathol, 182: 1–8.

Cross, S.S., McDonagh, A.J.G., Stephenson, T.J, et al. (1995) Fractal and integer-dimensional analysis of pigmented skin lesions. Am. J. Dermatol, 17:374–378.

Cutting, J.E. and Garvin, J.J. (1987) Fractal curves and complexity. Percept Psychophys, 42: 365–370.

D'Anselmi, F., Valerio, M., Cucina, A., Galli, L., Proietti, S., Dinicola, S., Pasqualato, A., Manetti, C., Ricci, G., Giuliani, A., Bizzarri, M., 2010. Metabolism and cell shape in cancer: a fractal analysis. Int. J. Biochem. Cell. Biol. 43, 1052e1058.

Dey, P., Mohanty, S.K. (2003). Fractal dimensions of breast lesions on cytology smears. Diagn Cytopathol 29(2):85Y86.

Don, S., Revathy, K. (2007). Classifying Mammogram Images using Fractal Features, Proc of IEEE ICCIMA, pp-539- 543.

Gazit, Y., Berk, D.A., Leunig M, et al. (1995) Scale-invariant behavior and vascular network formation in normal and tumour tissue. Phys Rev Lett; 75: 2428–2431.

Hlawatsch, A., Teifke, A., Schmidt, M., Thelen M. (2002) Preoperative assessment of breast cancer: sonography versus MR imaging. AJR Am J Roentgenol 179: 1493–1501.

Iftekharuddin, K.M, Jia, W., and Marsh, R. (2003) Fractal analysis of tumor in brain MR images. Machine Vision and Applications, 13:352–362.

Kerlikowske, K., Grady, D., Barclay, J. et al. (1996) Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA; 276: 33–38.

Kolb, T.M., Lichy, J., Newhouse, J.H. (2002) Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology 225: 165–175.

Landini, G. (2002). Fractals in Biology and Medicine. (Birkhauser Verlag, Basel).

Li, H., Liu, K.J., and Lo, S.C. (1997). Fractal modeling and Segmentation for the Enhancement of Microcalcifications in Digital Mammograms, IEEE Transactions on Medical Imaging, vol. 16, no. 6, pp. 785-798, Dec.

Losa G.A., Merloini, D., Nonnenmacher, T.F., Weibel, E.R. (2002) Fractals in biology and medicine. 1st ed. Basel: Birkhauser Verlag AG.

Luciani, M.L., Pediconi, F., Telesca, M.F. et al. (2011) Incidental enhancing lesions found on preoperative breast MRI: management and role of second-look ultrasound. La Radiologia Medica 116: 886-904, ISSN: 0033-8362.

Mandelbrot, B.B. (1975) Stochastic models for the Earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands. Proc. Natl. Acad. Sci. USA, 72: 3825–3828.

Margaret, T. et al. (2000) Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst; 92: 1081–1087.

Matsubara, T., Fujita, H., Kasai, S., Goto, M., Tani, Y., Hara, T. and Endo, T. (1997). Development of new schemes for detection and analysis of mammographic masses. In Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS97), pp 63Y66, Grand Bahama Island, Bahamas, December.

Michaelson, J.S., Cheongsiatmoy, J.A., Dewey, F., et al. (2005) Spread of human cancer cells occurs with probabilities indicative of a nongenetic mechanism. Br J Cancer; 93: 1244–1249.

Moon, W.K., Noh, D.Y., Im, J.G. (2002) Multifocal, multicentric, and contralateral breast cancers: bilateral whole-breast US in the preoperative evaluation of patients. Radiology 224: 569–576.

Mudigonda, N.R., Rangayyan, R.M., Desautels, JEL (2000) Gradient and texture analysis for the classification of mammographic masses. IEEE Trans Med Imag 19(10):1032-1043.

Mudigonda, N.R., Rangayyan, R.M., Desautels, JEL (2001). Detection of breast masses in mammograms by density slicing and texture flow-field analysis. IEEE Trans Med Imag 20(12):1215-1227.

Nguyen, T.M., Rangayyan, R.M. (2005). Shape Analysis of Breast Masses in Mammograms via the Fractal Dimension. In Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (CD-ROM), p 4, paper number 1852, Shanghai, China, September.

Normant, F. and Tricot, C. (1991) Methods for evaluating the fractal dimension of curves using convex hulls. Phys Rev A, 43: 6518–6525.

Norton, L. (2005). Conceptual and Practical Implications of Breast Tissue Geometry: Toward a More Effective, Less Toxic Therapy. Oncologist. 10(6):370-81.

Ohri, S., Dey, P., Nijhawan, R. (2004). Fractal dimension in aspiration cytology smears of breast and cervical lesions. Anal Quant Cytol Histol 26:109–112.

Pasqualato, A., Lei, V., Cucina, A. et al., (2013) Shape in migration: Quantitative image analysis of migrating chemoresistant HCT-8 colon cancer cells. Cell Adhesion Migr, 7(5): 450-459.

Pasqualato, A., Palombo, A., Cucina, A. et al. (2012) Quantitative shape analysis of chemoresistant colon cancer cells: correlation between morphotype and phenotype. Experim Cell Res, 318: 835-846.

Pasqualato, A., Palombo, A., Cucina, A., Mariggiò, M.A., Galli, L., Passaro, D., Dinicola, S., Proietti, S., D'Anselmi, F., Coluccia, P., Bizzarri, M., 2012b. Quantitative shape analysis of chemoresistant colon cancer cells: correlation between morphotype and phenotype. Exp. Cell. Res. 318, 835e846.

Pediconi, F., Catalano, C., Roselli, A. (2009) The challenge of imaging dense breast parenchyma: is magnetic resonance mammography the technique of choice? A comparative study with x-ray mammography and whole-breast ultrasound. Invest Radiol, 44 (7):412-421.

Peitgen, H.O, Jurgens, H., Saupe, D. (2004) Chaos and Fractals: New Frontiers of Science New York, NY: Springer.

Peitgen, H.O., (2002) Proceedings of the 6th International Workshop on Digital Mammography, Bremen, Germany, Springer-Verlag (ed.).

Pohlman, S., Powell, K.A., Obuchowski, N.A., Chilcote, W.A., Grundfest-Broniatowski, S. (1996). Quantitative classification of breast tumors in digitized mammograms. Med Phys 23(8):1337Y1345.

Rangayyan, R.M., Nguyen, T.M. (2007). Fractal Analysis of Contours of Breast Masses in MammogramsJ Digit Imaging. 20(3): 223–237.

Rangayyan, R.M., El-Faramawy, N.M., Desautels, J.E.L. (2000). Boundary modelling and shape analysis methods for classification of mammographic masses. Med Biol Eng Comput 38:487Y496.

Rangayyan, R.M., El-Faramawy, N.M., Desautels, J.E.L., Alim, O.A. (1997). Measures of acutance and shape for classification of breast tumors. IEEE Trans Med Imag 16(6):799Y810.

Rangayyan, R.M., Prajna, S., Ayres, F.J., Desautels, J.E.L. (2008) Detection of architectural distortion in prior screening mammograms using Gabor filters, phase portraits, fractal dimension, and texture analysis. Int J CARS 2:347–361.

Rosai, J. (2001). The continuing role of morphology in the molecular age. Modern Path, 14: 258-260.

Sahiner, B.S., Chan, H.P., Petrick, N., et al. (2001) Improvement of mammographic mass characterization using spiculation measures and morphological features. Med Phys 28(7):1455-1465.

Sahiner, B.S., Chan, H.P., Petrick, N., Helvie, M.A., and Hadjiiski, L.M. (2001). Improvement of mammographic mass characterization using speculation measures and morphological features. Medical Physics, 28(7):1455–1465.

Sankar, D., Thomas, T. (2009). Breast Cancer Detection using Entropy based Fractal Modeling of Mammogram. Int J Recent Trends Eng., 1(3) : 171-175.

Shelhamer M. (2007) Non-linear dynamics in Physiology. World Scientific, London.

Sickles, E.A. (1986) Mammographic features of 300 consecutive nonpalpable breast cancers. AJR Am J Roentgenol 146:661–663.

Smith, T.G., Lange, G.D., and Marks, W.B. (1996) Fractal methods and results in cellular morphology – dimensions, lacunarity and multifractals. J Neurosci.Methods, 69: 123–136.

Tourassi, G.D., Delong, D.M., Floyd, C.E. Jr. (2006) A study on the computerized fractal analysis of architectural distortion in screening mammograms. Phys Med Biol 51:1299–1312.

Velanovich, V. (1998). Fractal analysis of mammographic lesions: a prospective, blinded trial. Breast Cancer Res Treat. 49(3):245-9.

RESEARCH WITH X-RAY PERFORMED AT CIB (FOLLOWING CISB) FROM 1980 TO 1990

Roberto Cesareo(+), Giovanni E. Gigante(*)
(+) Istituto di Matematica e Fisica, Università di Sassari
(*) Dipartimento di Scienze di Base ed Applicate all'Ingegneria

Keywords: X rays, Tomography, Nuclear Analytical Techniques, X ray Fluorescence, Bone Mineral Content

Abstract: The activities carried out at the CIB in the decade of the eighties focused on some typical aspects at the time of the applications of nuclear techniques in biomedical areas such as the development of tomographic techniques, the analysis of biological samples and the development of techniques for measuring in vivo. We discuss some of the results obtained and also shown the dissemination activities carried out with the organization of international scientific conferences and the publication of books in English.

1 INTRODUCTION

The Center for Biomedical Engineering (CIB) was wanted by prof. Antonio Ruberti, at that time rector of the "La Sapienza", to develop research in the emerging field of bioengineering. Joined in this initiative from the beginning the group Nuclear Engineering, which even then was trying to diversify their own research, starting from knowledge, especially of an instrumental nature, which at that time were very advanced. Their choice fell on the techniques that were used X-rays in the previous decade had led to the development of the techniques of computed tomography, we can safely say, was the real beginning of the enormous development of diagnostic techniques in vivo. The Medical Physics was just beginning, the journal Medical Physics begins publications in the US in 1974, and bioengineering as well. It should be stressed that in 1977 Rosalyn Yalow (Nuclear Physicist) for the devopment of Radio Immune Assay (RIA) and in 1979 Hounsfield (Engineer) and Cormack (Nuclear Physicist) for the Computed Tomography received the Nobel Prize for medicine and physiology.
Initially the interest was fixed on nuclear analytical techniques that were used detection techniques originals that were based on the detection of a single photon (which at that time was a real innovation)

in order to develop techniques that could be automated. Later were developed simple prototypes of CT scanners for specialist applications including non-medical.
The great penetrating power of X-rays have, from the beginning, addressed their use in the investigations in which would be useful to analyze the entire sample to extract analytical information noninvasively. The good knowledge of the mechanisms of interaction of X-rays with matter and the very accurate knowledge of the absorption coefficients were allowed to develop different techniques.
The X-rays interact with matter by means of different effects, which can all be used in order to deduce information on the object of the interaction itself. In particular, the X-rays may (Figure 1):

- Cross, at least partially, the object (radiography and tomography in transmission)
- Give rise to diffuse X-ray Compton effect (identification of empty areas, Compton tomography)
- Give rise to secondary radiation X following a photoelectric effect (X-ray fluorescence analysis).

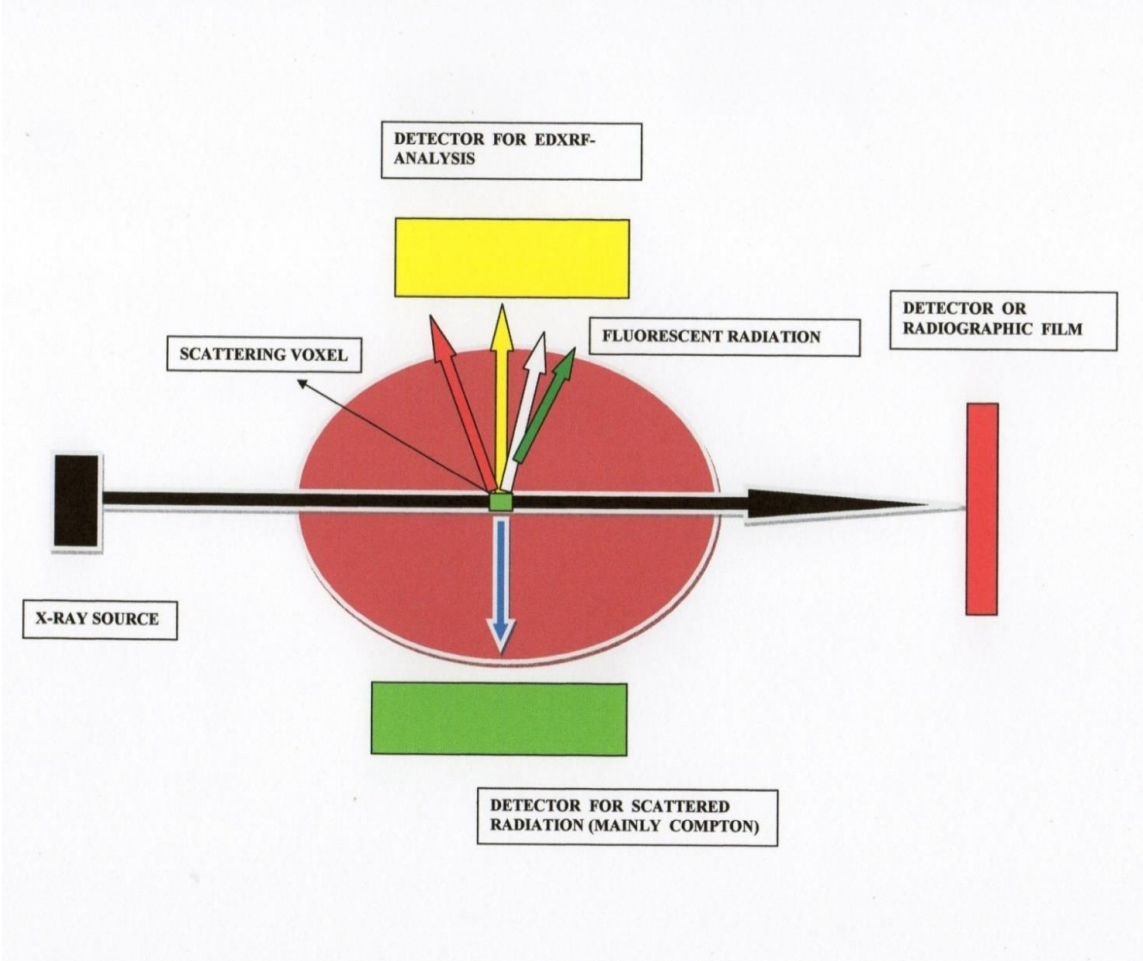


Figure 1 - X-rays emitted by a source (radioisotopic or tube X) may interact with a sample primarily through three effects: photoelectric effect, which has as consequence the emission of characteristic X-rays of the chemical elements present in the sample, the Compton effect, which corresponds to inelastic scattering of incident radiation, Rayleigh effect, which corresponds to the elastic scattering of the incident radiation. The three effects contribute to the attenuation of the incident X-beam.

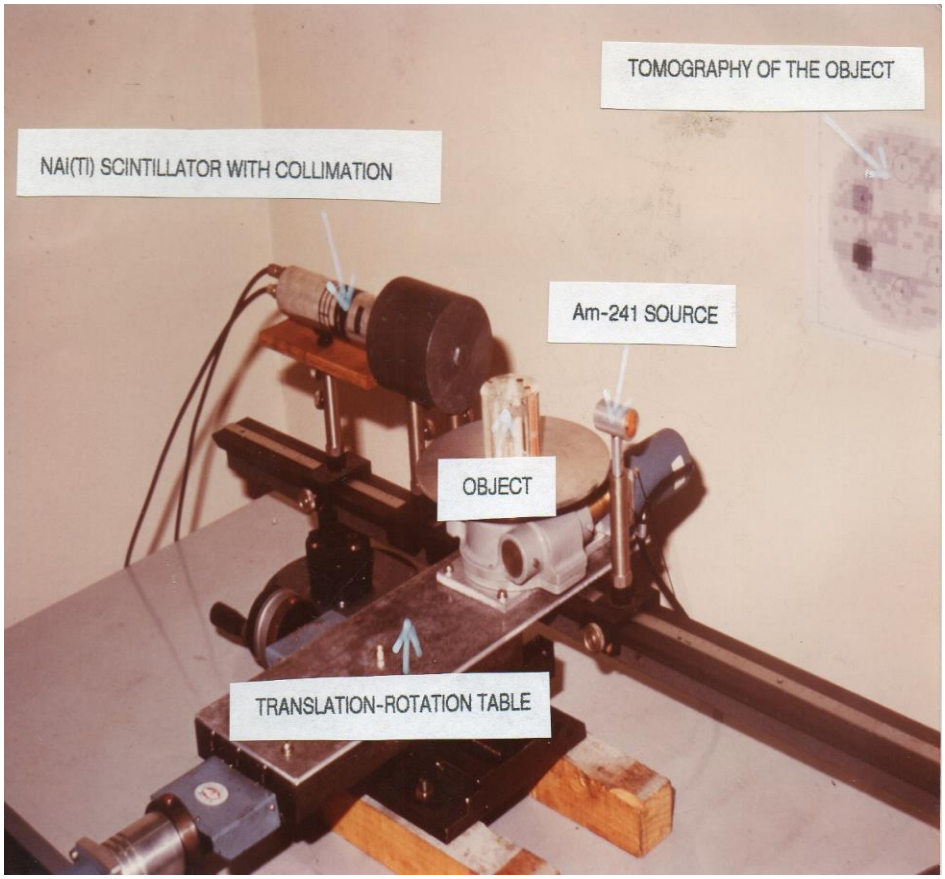


Figure 2 – Mini- X ray tomograph. They are visible: a. source of Am-241; b. detector NaI (Tl) collimated; c. a rotation-translation table on which there is the object whose tomographic image is shown in Figure 3.

2 Tomographic measures

In the early 80 has been designed and implemented a tomograph X (Figure 2), characterized by the following components (Cesareo, Giannini 1980, Cesareo, Giannini, Storelli 1982):

- radioisotope Am-241source;
-a single NaI (Tl) scintillator detector;
-a translation-rotation table;
-PC Apple for data processing and image reconstruction.

Tomography of a typical test object is shown in Figure 3. The prototype described above was improved over time, replacing the source of Am-241 with an X-ray tube, and the single detector with an image intensifier. They were also evaluated alternative methods, using the so-called "differential tomography" (Cesareo, Pedraza 1985, Cesareo1988, Cesareo, Mahtaboally 1989, R. Cesareo et al. 1989). An advanced version of the prototype described above has led to a model that has been marketed by the Company Gilardoni (Figure 4). It is particularly applied tomography to the study of soil physics (Crestana, Cesareo, Mascarenhas 1986, Cruvinel et al. 1990, Cesareo et al. 1988).

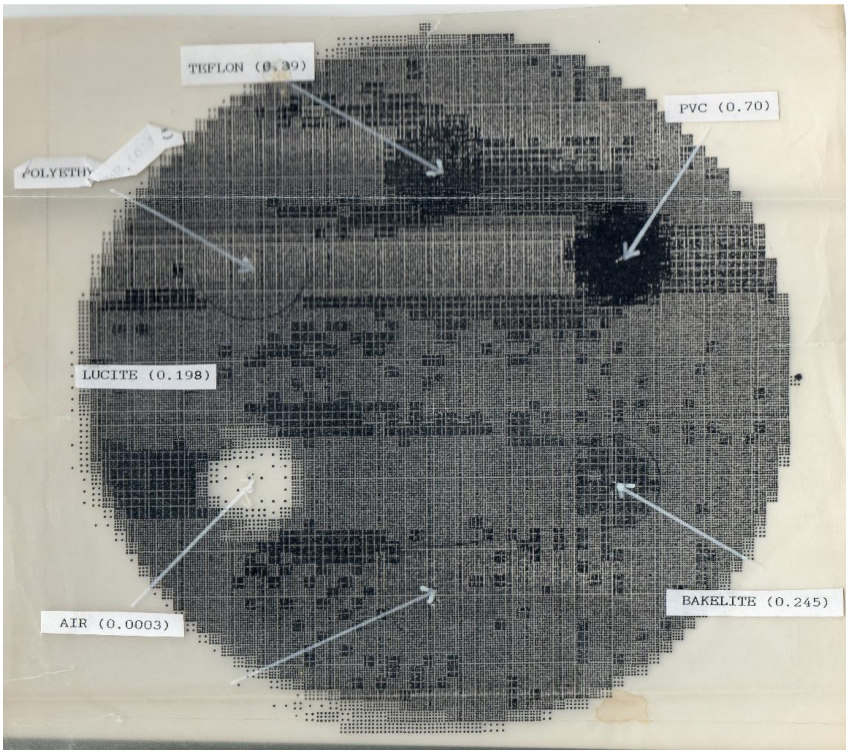


Figure 3 – Example of a tomography realized with the apparatus of Figure 2. The object imaged is a Plexiglas cylinder with six cylinders of attenuation material similar to that of Plexiglas (PVC with linear attenuation coefficient = 0.7; Bakelite, linear attenuation coefficient = 0.245; Lucite with linear attenuation coefficient = 0.198; air with linear attenuation coefficient = 0.0003; polyethylene with linear attenuation coefficient = 0.185; Plexiglas with linear attenuation coefficient = 0:19.

Figure 4 – Tomograpic X apparatus marketed by Gilardoni of Mandello del Lario

3 Assessment of Bone Mineral Content (BMC) with scattered radiation

Following the approach given in the introduction, research also developed in the direction of extracting information of clinical interest from the scattered radiation.

In the mid-eighties were conducted systematic research on the use of coherent and incoherent radiation for non-destructive analysis of materials. In the 1986 appear a paper in Medical Physics describing the possibility to assess the BMC using the ratio of coherent and incoherent radiation (R/C ratio) (Gigante, Sciuti 1986). The results obtained were quite good, but in the following, the development of new apparatus for measurements of BCM using transmitted radiation and quasi-monochromatic radiation from an X-ray tube had such advantages as to make the use of scattered radiation useless. In figure 5 is shown the apparatus developed at the CIB for the analysis of BMC.

The idea on which is based this method was that, despite being heavily penalized coherent radiation at large scattering angles geometry, which allowed to achieve a system of investigation very compact would have favored the clinical use of the system. The ability to simultaneously measure the coherent and incoherent scattered radiation it is possible to make a more accurate measure of the BMC.

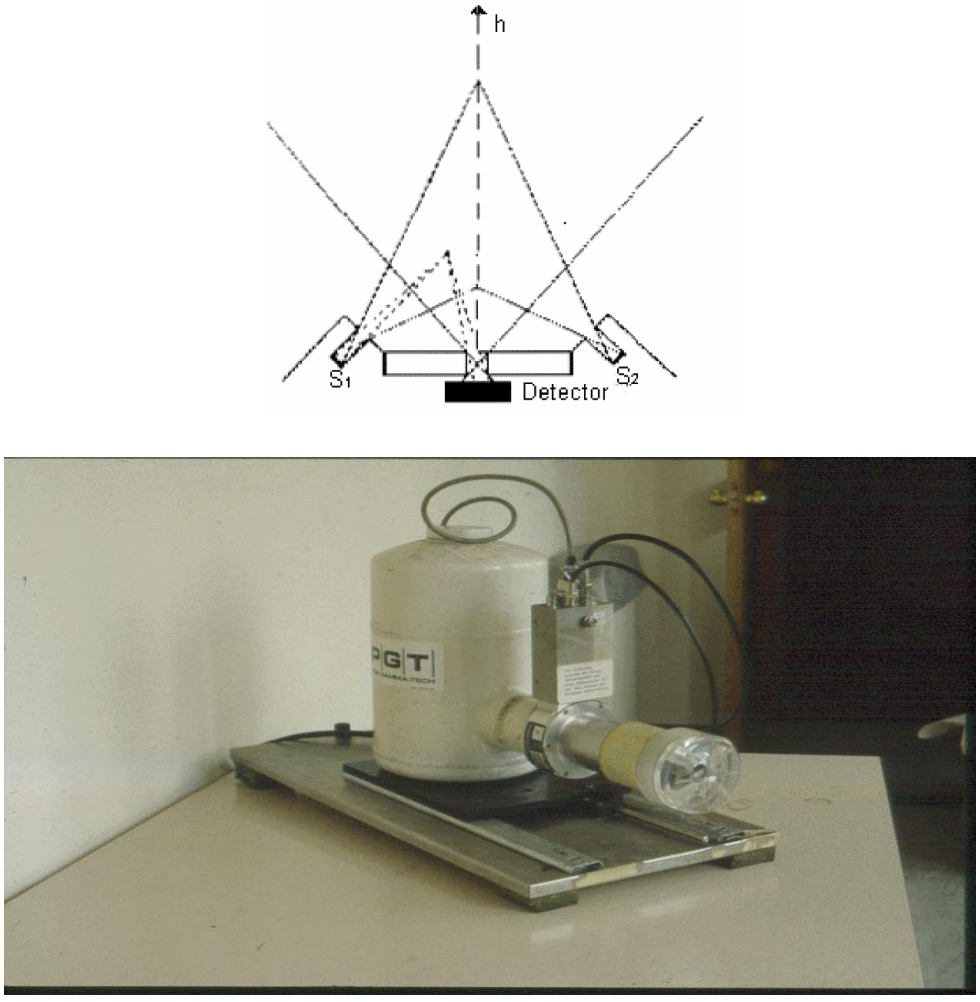


Figure 5 – Schematic drawing of a two source detector assembly (upper) and prototype spectrometer for the measurement of BMC

4 XRF Analysis

X ray Fluorescence (XRF) is an analytical technique developed in the early fifties as a powerful tool for multielemental laboratory analysis of many samples (rocks, ground, living tissues, industrial products, etc.) using the technology of wavelength dispersive detectors. The success of this technique is undisputable in many area of science and technology. Starting from the 1960's and mostly in the 1970s the single photon counting techniques were developed, especially in the nuclear physics, to detect high energy particles (and in particular x-rays). This innovation has allowed the development of very compact and easy to use equipment that in a few years have become an indispensable tool in non-destructive testing. At first there was a considerable attention for the use of such systems in the medical field for both in vitro assays and in vivo

In the '80 at CIB has been developed a system for laboratory analysis of X-ray fluorescence of biological samples, in particular were analysed with satisfactory results the following type:

- blood and serum samples deposited on the filter (Cesareo, Viezzoli 1983);
- hair samples (Cesareo 1984);
- samples of kidney stones (Cesareo 1985);
- stable tracers;
- gunshot residues (Cesareo et al. 1988).

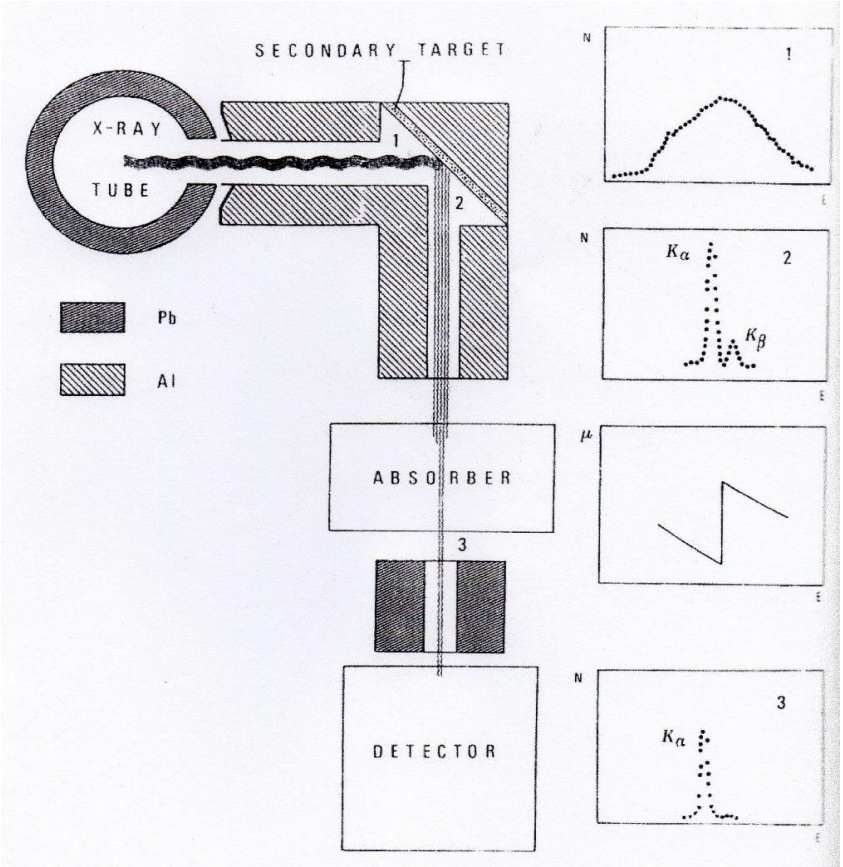


Figure 6 – System for mono-chromatization of an X-rays beam. The X-rays, passed through a collimator, impinge on target consisting of a single element (e.g. Ag) causing photoelectric effect with the emission characteristics of the lines of this element. These characteristic lines (see spectrum 2) are made to cross a second element of adequate thickness and photoelectric discontinuity which has energy between the K_α and K_β lines of the first element. The end result is the K_α line of "collimator" element (see spectrum # 3 in the lower right).

The apparatus was constituted by an X-ray tube by 80 kV and 40 mA and an X-ray detector in Ge-high purity. The emitted radiation from the tube was monochromatised through the use of secondary target (Figure 6), in order to optimize the radiation energy to the elements to be analysed. A typical X-ray spectrum of a sample of 0.5 ml of blood serum collected on filter paper bibulous is shown in Figure 7 (Cesareo 1988).

In the 1982 was held at CIB a international Workshop on "X-ray Fluorescence (XRF and PIXE) in Medicine" whose proceedings were published in a book edited by prof. Cesareo and printed by Acta Medica (fig.8). In the 1989 the second edition of this workshop held in Capri with the participation of CIB the proceeding were published with the collaboration of prof. Cesareo.

5 Conclusions

This short report on the activities of the CIB in the eighties is to show how this center has been very active from the beginning in the sectors that from time to time have alternated in landscape of biomedicine. In particular, has been showed that from the beginning there was a strong emphasis on the development of methods that have a good fallout in the medical field.

The subsequent transformation of the CIB in CISB shifted attention to other research subjects, equally significant.

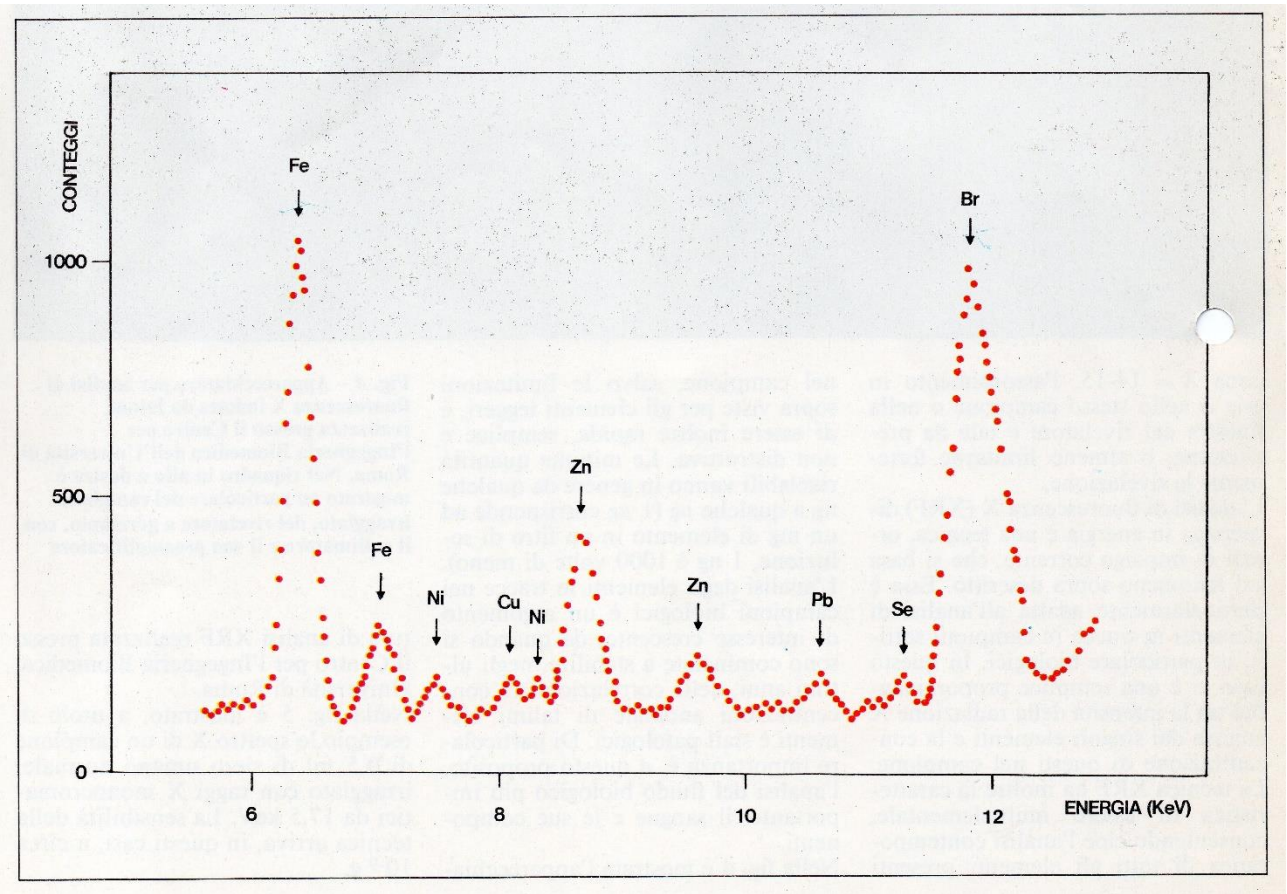


Figure 7 – X fluorescence spectrum of a sample of 0.5 ml of serum deposited on a filter paper bibulous. The measured concentrations are (in ppm): Fe = 1.82, Ni = 12:18, Cu = 0.2, Zn = 0.65, Pb = 0.09 Se = 0.09 Br = 1.18).

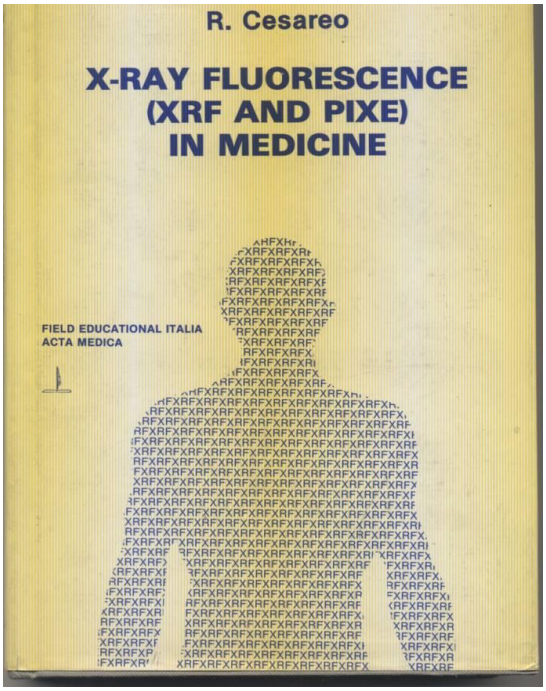


Figure 8 – Frontispiece of proceedings of 1985 Rome conference.

References

Cesareo R., M. Giannini: Elemental analysis by means of X-ray attenuation measurements; Nucl. Meas. Methods 169 (1980) 551.

Cesareo R., Giannini M., Storelli L.: A miniature X-ray tomography scanner employing radioisotope sources; Int. Conf. on applications of physics to Medicine ; Trieste, 30.3-3.4.1982; In Proc. of the Conference p. 631.

Cesareo R., Pedraza L.J.: Differential attenuation of X-rays: analytical applications; Nucl. Instrum. Methods A239 (1985) 367.

Cesareo R.: Principles and applications of differential tomography; Nucl. Instrum. Methods A270 (1988) 572.

Cesareo R.,. Mahtaboally S.Q.G: Use of differential tomography in the study of natural processes; IEEE Trans. Med. Imaging 8 (1989) 163.

Cesareo R., Mahtaboally S.Q.G., Lal U.S., Storelli L., Castellano A., Bonpadre S., Ferrante L., Leone L., Rescigno A., Salleo S.: Applications of differential tomography to biological systems; Phys. Med. 1 (1989) 17.

Crestana S., Cesareo R., Mascarenhas S. : Using a computed tomography miniscanner in soil science; Soil Sci. 142 (1986) 56.

Cruvinel P.E., Cesareo R., Crestana S., Mascarenhas S. : X-ray and gamma-ray computerized monitomograph scanner for soil science; IEEE Trans. on Instrum. & Measurem., New York 39 (1990) 745.

Cesareo R., S. Crestana, P.E. Cruvinel, Mascarenhas S. : The mini-CT scanner: physical principles and applications in soil physics; Escuela latino-americana de fisica de suelos 24.1 – 6.2.1988.

Cesareo R., Mascarenhas S. : A new tomographic device based on the detection of fluorescent X-rays; Nucl. Instrum. Methods A277 (1989) 669.

Gigante G.E. and Sciuti S. "A large angle Coherent/ Compton scattering method for measurement "in vitro" of trabecular bone mineral concentration" Medical Physics,12(3), 321-326, 1985.

Cesareo R. and Viezzoli G.: Trace element analysis in biological samples by using XRF spectrometry with secondary radiation; Phys. Med. Biol. 28 (1983) 1209.

Cesareo R.: Nuclear analytical techniques in medicine; Proc. II int. Conf. appl. of phys. to medicine & biology; World Sci. Publ. , Singapore (1984) 219-235.

Cesareo R.: Biomedical applications of energy dispersive X-ray fluorescence; Trends in anal. Chemistry 4 (1985) 65.

Cesareo R., Alexiu E., Sacchetti G., Merli S..: EDXRF for gunshot residue analysis; Phys. Med. 2 (1986) 71.

Cesareo R., Photon induced X-ray emission in Nuclear analytical techniques in medicine; Ed. R. Cesareo, Elsevier, Amsterdam, 1988.

PERSONAL EXPOSURE TO SOLAR ULTRAVIOLET RADIATION: WHY AND HOW TO ASSESS IT

Giuseppe Rocco Casale
Department of Physics Sapienza University of Rome, Piazzale Aldo Moro 2, 00185 Rome, Italy
giuseppe.casale@uniroma1.it

Anna Maria Siani
Department of Physics Sapienza University of Rome, Piazzale Aldo Moro 2, 00185 Rome, Italy
annamaria.siani@uniroma1.it

Alfredo Colosimo
Department SAIMLAL Sapienza University of Rome, Piazzale Aldo Moro 2, 00185, Rome, Italy
alfredo.colosimo@uniroma1.it

Keywords: Solar radiation, ultraviolet, personal exposure, dosimetry, polysulphone.

Abstract: Solar ultraviolet (UV) radiation has deeply influenced the evolution of life on earth and it was among the possible drivers of different skin pigmentation in humans: at low latitudes, where solar radiation is on average intense all the year, darker skins help to protect from the deleterious effects of radiation, while at higher latitudes fair skins help to maximize vitamin D production from much lower ambient UV. Changes in habits and attitudes in recent years combined with the increase in human migration from original areas have made human skin pigmentation no longer necessarily suited to the environment. Furthermore, the documented stratospheric ozone downward trend due to chlorofluorocarbons and its slow recovery have been associated, at least theoretically, with an increase of solar UV at the earth's surface. In this contribution we want to show the importance of assessing solar UV personal exposure by means of dosimetry, particularly focusing on a dosimetric technique based on the polysulphone polymer. Some results of dosimetric field campaigns, performed with groups of volunteers selected among the Italian population, are also presented.

1 INTRODUCTION

The spectrum of solar radiation outside the atmosphere can be approximated by that of a black body at the temperature of about 5800 K. Approximately 9% of the extraterrestrial solar radiation is in the ultraviolet (UV), divided into three bands of different wavelength: UVA (400-320 nm), UVB (320-290 nm) and UVC (290-200 nm). Penetrating the atmosphere, approximately 50% of the total UV radiation is lost: UVC is totally absorbed by ozone (O_3) and oxygen (O_2) in the upper atmosphere (above 50 km) and has no effect on human health, more than 90% of UVB is absorbed by atmospheric O_3 in the stratosphere (mainly in the region between 20 and 25 km) while UVA reaches the surface almost unaffected. Both UVB and UVA carry energy of the order of 1 eV but the first spectral range is more active than the latter in determining biological effects on humans.

A global view of the solar UV reaching the earth's surface for the year 2013 is given in Figure 1, where the yearly average of the daily dose of solar UV radiation, weighted with the erythemal action spectrum, is presented. The different colors from purple (low UV) to yellow-green (high UV) evidence that the expected latitudinal gradient, due to the decreasing sun flux from the equator towards the poles, is not the solely effect. Peak values can indeed be observed over high mountain ranges, e.g. Andes and Himalaya (Cordero et al., 2014), while the breaks in the longitudinal simmetry can be attributed to sky cloudiness and surface albedo. It can be summarized that solar UV radiation on a horizontal plane[1] depends on:

- solar zenith angle SZA, i.e. the angle between the sun and the local vertical;
- atmospheric attenuation, mostly by O_3 and aerosols;
- clouds, which attenuate solar UV primarily by scattering;
- surface albedo, since most ground surfaces reflect UV radiation less than 10% but snow can reach up to 90%;
- altitude, since higher altitudes mean reduced atmospheric optical thickness.

Irradiance is measured by ground-level instruments as spectroradiometers, narrow- and broad-band monitors or retrieved by satellite measurements. In Italy an official network for solar UV measures does not exist but the Solar Radiometry Observatory of Sapienza University of Rome (41.9°N, 12.5°E, 75 m asl) regularly measures ground-based spectral UV irradiance since 1992. The stress due to solar UV radiation during the evolution of life on earth is likely one of the main responsibles for the development of different skin pigmentation in humans (Juzeniene et al., 2011). Generally speaking, the inhabitants at low latitudes have darker skin pigmentation also to protect themselves from the deleterious effects of UV while those in higher latitudes have fair skin to maximize the beneficial effect of vitamin D production (McKenzie et al., 2009). However, UV is not the only factor to determine the different pigmentation of human beings, as shown in Figure 2 by the global distribution of skin color for native populations before 1940: additional factors, as diet and sexual selection, must be taken into account to explain the non exact overlap between Figure 1 and Figure 2.

In the recent years there has been an increase in human migration over all directions (south-north and east-west) and an updated version of Figure 2 would be much more difficult to obtain. One of the consequences of such huge movements is that skin pigmentation is no longer necessarily adapted to the environment. At low latitudes, dark-skinned populations have very low incidence of melanoma and skin cancers but they suffer for the incidence of rickets and osteomalacia when living at high latitudes (El Ghisassi et al., 2009). At the same time, the incidence of melanoma and non-melanoma skin cancers has dramatically increased in fair-skinned populations, with a clear contribution from changing habits and clothing, modified postures in the sun, increased longevity and worldwide ozone depletion (WHO, 2010).

Ambient exposure (AE) is determined in two steps. The first consists in weighting solar UV irradiance $I(\lambda)$ at the surface, dependent on the wavelength λ, with a response function (e.g. erythema (CIE, 1998)) $S(\lambda)$ called action spectrum, a weighting function of the effectiveness of UV radiation of different wavelengths in inducing a given biological effect. Examples of action spectra for erythema, vitamin D production (CIE, 2006) and human non melanoma skin cancer (de Gruijl and van der Leun, 1994) are shown in Figure 3.

[1] The amount of energy coming from the sun and the sky and reaching the unit of terrestrial horizontal surface in the unit of time is named *irradiance*.

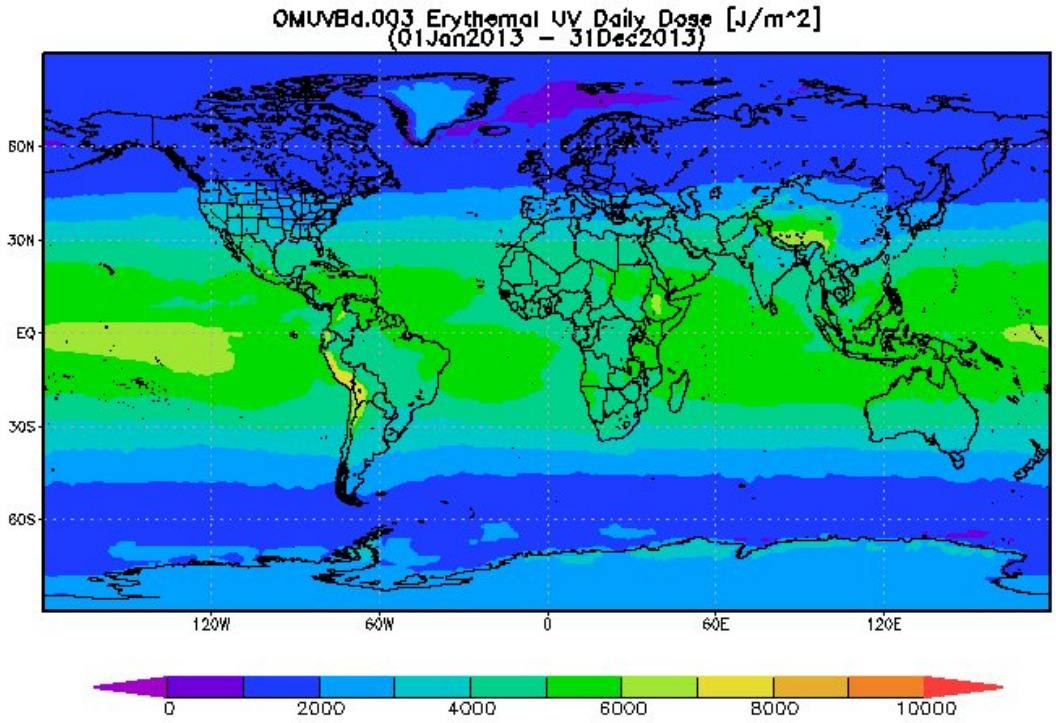


Figure 1: Global erythemally weighted solar UV daily dose in 2013 (source: "Giovanni" online data system, developed and maintained by the NASA GES DISC).

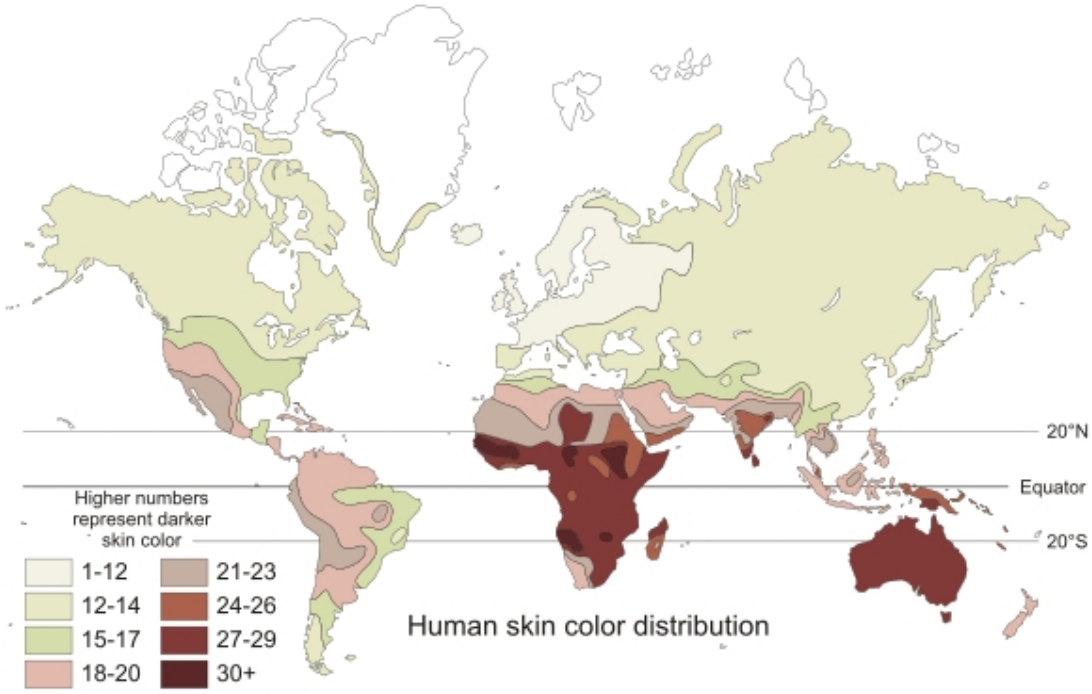


Figure 2: Geographic distribution of skin colors for native populations prior to 1940 according to the Von Luschan's chromatic scale (source: en.wikipedia.org).

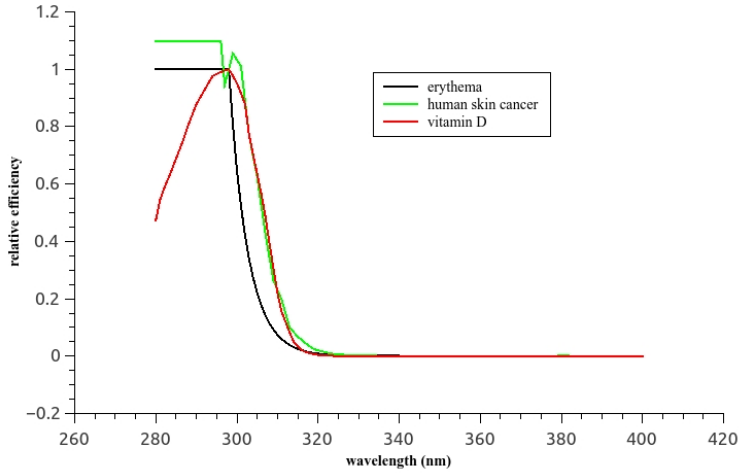


Figure 3: Examples of action spectra for erythema (black), vitamin D production (red) and human skin cancer (green) normalized to unity at 298 nm.

Action spectra, expressed in relative units, represent the efficacy of radiation in initiating a given biological effect. Human action spectra are derived following different approaches: e.g. the erythemal action spectrum in Figure 3 (black line) is an idealized spectrum based on the average of experimental spectra on humans, each obtained using multiple monochromatic light sources; the human carcinogenesis spectrum (green line) is indeed generated in mice exposed to filtered UV radiation from xenon arc lamps but transferred to humans after complicated calculations taking into account differences in epidermal thickness. By integrating the spectral convolution of irradiance and action spectrum within an arbitrarily chosen wavelength interval $[\lambda_1,\lambda_2]$, it is possible to determine the dose rate T, that is the time dependent UV energy able to produce a detectable biological effect in that spectral range:

$$T(t) = \int_{\lambda_1}^{\lambda_2} I(\lambda) S(\lambda)\, \mathrm{d}\lambda \tag{1}$$

The second step implies the integration of the dose rate T over an arbitrary time interval $[t_1,t_2]$ (hour, day, month, year or other), so that AE is determined:

$$AE = \int_{t_1}^{t_2} T(t)\, \mathrm{d}t \tag{2}$$

Although UV measuring instruments have reached a high level of accuracy and action spectra are more and more well defined, ambient exposure does not accurately represent the population distribution of individual exposures to the sun (Casale et al., 2009). One reason is that the geometrical difference between a horizontally fixed detector and the curved body surface produces significant deviations in exposure. In addition, different postures, activities and attitudes among individuals lead to hundred-fold differences in personal solar UV exposure in the same environment (Siani et al., 2014). Finally, biological variations in skin pigmentation and the use of sun protective devices, such as sunscreens or hats, are elements of additional complexity (Casale et al., 2009). What we know for sure is that personal exposure[2] to the sun can be both beneficial and dangerous, depending on the dose accumulated by each subject. Table 1 is a summary of the known effects of solar radiation, mostly due to the UV portion of the spectrum but also to the visible range (VIS: 380–700 nm).

[2]The *dose rate* T is also referred to as *exposure dose rate* and the *dose* as *exposure dose* or simply *exposure*. The UV dose is the *radiant exposure* accumulated individually on the external surface or interface and not the *absorbed dose*, unlike its counterpart in medicine where it refers to the dose of a chemical agent that is actually absorbed by a target tissue.

Table 1: Summary of solar radiation effects on human beings as a function of time (Juzeniene et al., 2011). For each effect the most effective spectral range is indicated.

Immediate	**Delayed**	**Chronic/Latent**
Vitamin D synthesis (UVB)	Erythema (sunburn) (UVB)	Photoageing (UVB+UVA)
Immediate pigmentation (UVB+UVA)	Delayed pigmentation (UVB)	Photocarcinogenesis (UVB+UVA)
Photoimmunosuppression (UVB)		Catataract (UVB+UVA)
Nitric oxide release (UVA+VIS)		Photokeratitis (UVB)
Blood pressure decrease (UVA+VIS)		
Antidepressant (VIS)		
Folate photodegradation (UVB)		

2 POLYSULPHONE DOSIMETRY

Solar UV dosimetry comprises a group of sensors and methodologies able to evaluate the exposure of differently oriented body surfaces. UV dosimeters can be based on physical, biological or chemical effects. Physical badges are miniaturized electronic dosimeters that can be considered the best technique to quantify personal exposure mainly over long periods (Seckmeyer et al., 2012) but their high cost limits the use at large scale. Biological dosimeters can be also used for long term measurements but the associated technique involves a sophisticated time-consuming calibration and readout process. Chemical dosimeters, as polysulphone (PS) badges, have the main limitation in their reduced dynamic range as they are saturated after relatively short exposure but on the other side they are hygro-thermally stable, portable and cheap (Siani et al., 2014). PS dosimeters have been used for over 30 years to assess personal solar erythemally effective UV exposures, since the PS spectral response resembles that of erythema (Figure 4), and they were also adopted by the authors for their dosimetric studies.

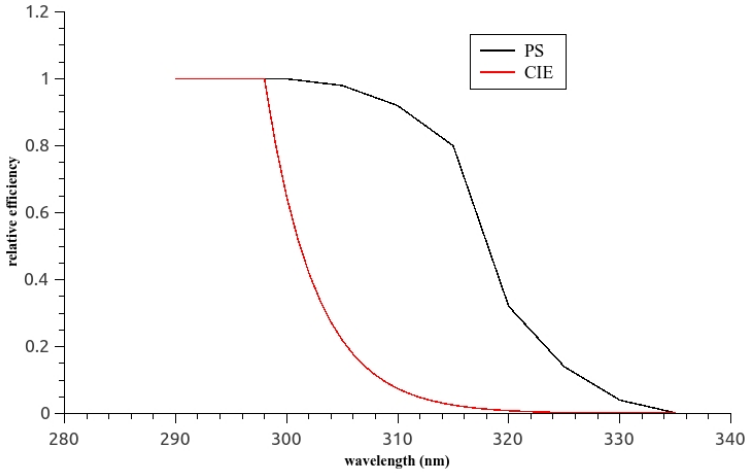


Figure 4: Comparison between the erythemal action spectrum (red) and the PS response (black) between 290 and 335 nm. Both spectra are defined by CIE and normalized at 298 nm.

PS is a polymer which undergoes a photo-degradation upon exposure to UV radiation, generating a change of absorbance A (indicated as ΔA) with a peak at 330 nm. PS badges are made up with a film 40 μm thick, a right compromise between an acceptable spectral response with respect to the erythemal action spectrum and the ease of handling. Due to the intrinsic mismatch between the PS response and the erythema action spectrum, the determination of a calibration curve, i.e. the relationship between ambient exposures and the corresponding PS absorbance changes, is needed. PS dosimeters are calibrated outdoors by direct comparison with a ground-based instrument which provides the ambient exposure: they are exposed horizontally to the sun in an appropriate number (usually between 10 and 20) and removed at given time intervals. When exposure increases, dosimeters increase their absorbance till saturation.

The general form of the erythemal dose–response curve is represented by a third–degree polynomial fitting the erythemally effective ambient exposure and the corresponding film absorbance changes at 330 nm:

$$AE = c(\Delta A + \Delta A^2 + 9\Delta A^3) \tag{3}$$

It was outlined by the authors that caution should be used when PS horizontal calibrations are performed under high albedo conditions and/or high incidence angle of the sun's rays on the film, approximately 70° (Casale et al., 2012) Within such limitations, the horizontal calibration can be transferred to non horizontal dosimeters worn by humans on different body parts and equation (3) becomes:

$$D = c(\Delta A + \Delta A^2 + 9\Delta A^3) \tag{4}$$

where D is now the personal dose at a specific body site represented by the dosimeter whose absorbance change is ΔA. An example of calibration curve is reported in Figure 5 for dosimeters exposed on a horizontal plane at Sapienza University of Rome on 23 March 2012. The cubic fit is represented by the continuous thin line.

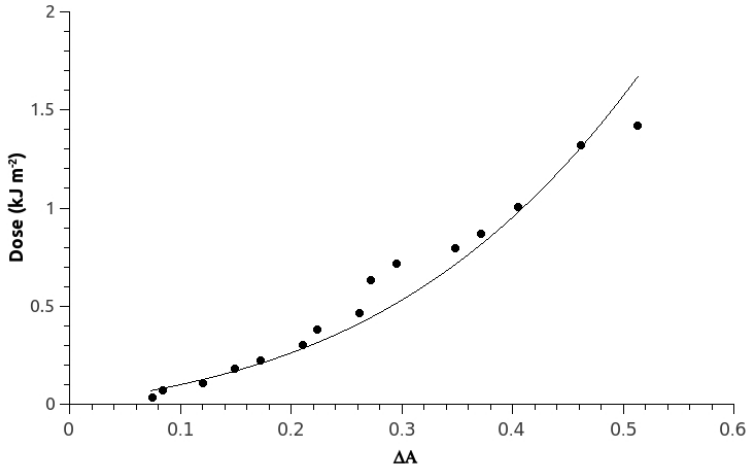


Figure 5: PS calibration curve for horizontally exposed dosimeters at Rome on 29 March 2012. Black dots are real measures, each point representing the change in absorbance (horizontal axis) due to an erythemally weighted dose (vertical axis) in kilojoules per square meter. The cubic fit (equation (4)) is shown as a black thin line, determined by least squares method. The multiplying coefficient is $c = (0.84 \pm 0.03) kJm^{-2}$ with explained variance R^2=0.94.

3 FIELD EXPERIMENTS WITH VOLUNTEERS

The authors have been organizing field campaigns in Italy with selected volunteers since 2004 with the aim of quantifying personal exposure during occupational or recreational activity. Each study was driven by dermatological or epidemiological issues and was stimulated by the great variability of the surface solar UV field over the Italian territory, characterized by complex orography and high latitude stretch. During the field campaigns, horizontal calibrations were performed covering the same temporal range of personal exposures and volunteers were asked to wear one or more PS dosimeters on a specific body site, depending on their occupation or leisure time activity. In addition, it was found useful to collect ancillary quantities in order to have as much information as possible to correlate with individual exposure: skin color (immediately before and after exposure) by means of skin reflectance spectrophotometry; phototype according to Fitzpatrick classification by visual inspection (Fitzpatrick, 1988); skin temperature (pre and post exposure) by means of radiometric thermometer; free radicals in the blood through blood analysis, performed some days before and following sun exposure; sex, age, sunscreen use, habits etc. through personal questionnaires.

Table 2 summarizes the field experiments with volunteers, performed by the authors, whose results were

already published: skiers (Siani et al., 2008), sunbathers (Siani et al., 2009) and vineyard growers (Siani et al., 2011). Details for each field experiment can be found in the papers. The last column of Table 2 shows the median ER, Exposure Ratio, representative for each category. ER is defined as the ratio between the personal exposure estimated by means of equation (4) and the corresponding ambient exposure measured by an instrument or retrieved from a horizontally located PS dosimeter. ER is dimensionless and gives the chance to compare results from different dosimetric techniques under different environmental conditions, provided that an accurate estimate of the related uncertainty is carried on. ER can be expressed as a decimal number or the corresponding percentage (e.g. 0.25 or 25%) and can be higher than 100% when the dose is higher than the ambient exposure (see Table 2).

Table 2: Summary of field experiments of personal exposure carried out by the authors by using PS dosimeters. Only published results are reported. The first column shows the dates of the campaign, one single day or more consecutive days. The second column reports the category (occupational or recreational) object of the study. The third column refers to the body site where the PS were attached as representative of the group going to be exposed while the last column is the median ER calculated over all available individual ERs.

Dates (year/month/day)	**Category**	**Body site**	**ER**
2006/03/30-2006/04/04	Skiers	Forehead	1.02
2007/01/29-2007/01/30	Skiers	Forehead	0.60
2005/05/27	Sunbathers	Chest	0.19
2005/04/20-2005/04/22	Vineyard growers	Back	0.72
”	”	Arm	0.44
2005/07/12-2005/07/13	”	Back	0.50
”	”	Arm	0.29
2005/10/11-2005/10/12	”	Back	1.00
”	”	Arm	0.67

The uncertainty that can be attributed to ER data is at least 20%, although other factors (such as the correct positioning of PS dosimeters on the chosen body site and their appropriate use for the entire exposure, not always directly inspected by the authors) could increase it. ER median values show a large variability, primarily due to the different posture of the individuals: higher ER values are reached when the dosimeter is almost horizontal (as in the case of the back of vineyard growers) but also if the dosimeter in a near vertical position (as for the skiers) is subject to elevated reflection caused by snow albedo. It is also to be kept in mind that low ER values do not necessarily mean low personal exposures, that are indeed retrieved multiplying ER by the corresponding ambient exposure. In any case, particular caution is needed with PS dosimeters when the horizontal calibration is transferred to tilted planes.

Several statistical approaches were adopted to find a relationship between ER and the ancillary parameters, that could directly or indirectly be affected by sun exposure. Although results are often not easy to interpret due to the high individual variability, it seems that some aggregation methods as cluster analysis can help. Furthemore, ancillary information could be more valuable when new studies will be published in the literature: a recent paper, for instance, focused on the potential role of skin temperature in skin cancers when coupled with sun exposure (Petersen et al., 2014).

4 CONCLUSIONS AND FUTURE WORK

It is of fundamental importance to quantify personal exposures to solar radiation, mainly in the UV portion of the spectrum, since many health effects are directly or indirectly linked to the amount of photons accumulated by target tissues (skin or eye: Table 1). For this reason, the research for new UV dosimeters has led to the newer electronic monitors based on AlGaN detectors having a spectral response that matches the erythema action spectrum with high accuracy. In addition, they record UV as a function of time and not just as integral over time (as non-electronic dosimeters do). On the other side, PS dosimeters have been used for solar UV exposure assessment for more than thirty years and they represent an important comparative tool for upcoming electronic badges.

In the present study, the authors have explained why and how to assess personal exposure to solar UV by using data from their work with PS dosimeters (Table 2): they show an acceptable response similar, although not very accurate, to the erythema action spectrum. The resulting large uncertainty on ER (20% or more) is indeed sufficient for most practical purposes and indeed new ER data for Mediterranean populations were provided. PS dosimeters are not expensive, making it possible to organize field campaigns with volunteers at a moderate cost, with statistically significant results. In addition they are light and portable, can be easily attached on the chosen body site and they are thermally stable with a known dark reaction. Their main problems reside in the fact that they reach saturation after some hours of exposure and cannot be reused. For such purpose, new chemical dosimeters based on PPO (poly-phenylene-oxide) and PVC (poly-vinyl-chloride) films with extended dynamic range have been developed and need to be tested. Recent findings from the authors also show that the field calibrations (Figure 5) of PS badgfes should be performed under the same environmental conditions of the exposed category and not extend for long time intervals, possibly symmetric with respect to the solar noon and with SZA below 65°-70°.

Results from field campaigns (Table 2) show an elevated variability in the Exposure Ratio, due both to the different activities and postures of individuals and to the non homogeneous radiative scenarios. It seems clear that dosimetry allows to determine differences in exposure levels which otherwise would not emerge. A methodology for personal exposure studies was proposed and applied combining personal doses with colorimetric, skin temperature and other ancillary information.

Recently the authors investigated the possibility to use PS film for other UV-dependent biological effects, besides erythema (Figure 4). They found that the PS response function shows theoretically a reasonable similarity with the action spectra of vitamin D production and human skin cancer (Figure 3) but also photocheratitis, ACGIH and photoimmunosuppression. These results need to be validated with ad hoc measurements but in principle they open the way to personal exposure studies in which multiple skin and eye effects could be simultaneously studied within a single field experiment provided that the dosimeters positioning is properly chosen and the calibration curve for each separate effect is determined.

REFERENCES

Bouillon, R., Eisman, J., Garabedian, M., Holick, M., Kleinschmidt, J., Suda, T., Terenetskaya, I., Webb, A. (2006). *Action spectrum for production of pre-vitamin D3 in human skin*. Commission Internationale de l'Eclairage (CIE), Technical Report 174, Central Bureau, Vienna, Austria.

Casale, G.R., Siani, A.M., Colosimo, A. (2009) Polysulphone dosimetry: a tool for personal exposure studies. *Biophysics & Bioengineering Letters*, 2(1):1 – 14.

Casale, G.R., Siani, A.M., Diémoz, H., Kimlin, M.G., Colosimo, A. (2012) Applicability of the Polysulphone Horizontal Calibration to Differently Inclined Dosimeters. *Photochemistry and Photobiology*, 88:207 – 214.

CIE (1998). *Erythema reference action spectrum and standard erythema dose*. Commission Internationale de l'Eclairage (CIE), Joint ISO/CIE Standard ISO 17166:1999/CIE S007-1998, Vienna

Cordero, R.R., Seckmeyer, G., Damiani, A., Riechelmann, S., Rayas, J., Labbec, F., Larozed, D. (2014) The world's highest levels of surface UV. *Photochemical and Photobiological Sciences*, 13:70 – 81.

de Gruijl, F.R. and van der Leun, J.C. (1994) Estimate of the wavelength dependency of ultraviolet cacinogenesis in humans and its relevance to the risk assessment of stratospheric ozone depletion. *Health Physics*, 67:319 – 325.

El Ghisassi, F., Baan, R., Straif, K., Grosse, Y., Secretan, B., Bouvard, V., Benbrahim-Tallaa, L., Guha,N., Freeman, C., Galichet, L., Cogliano, V. (2009) WHO International Agency for Research on Cancer Monograph Working Group: A review of human carcinogens-Part D: radiation. *Lancet Oncology*, 10:751-752.

Fitzpatrick, T.B. (1988) The validity and practicality of sun-reactive skin types I through VI. *Archives of Dermatology*, 124:869-871.

Juzeniene, A., Brekke, P., Dahlback, A., Andersson-Engels, S., Reichrath, J., Moan, K., Grant, M.. Moan, J. (2011) Solar radiation and human health. *Reports on Progress in Physics*, 74:066701.

McKenzie, R.L., Liley, J.B., Bjorn, L.O. (2009) UV Radiation: Balancing Risks and Benefits. *Photochemistry and Photobiology*, 85:88 – 98.

Petersen, B., Philipsen, P.A., Wulf, H.C. (2014) Skin temperature during sunbathing - relevance for skin cancer. *Photochemical and Photobiological Sciences*, 13(8):1123 – 1125.

Seckmeyer, G., Klingebiel, M., Riechelmann, S., Lohse, I., McKenzie, R.L., Liley, J.B., Allen, M.W., Siani, A.M., Casale, G.R. (2012) A Critical Assessment of Two Types of Personal UV Dosimeters. *Photochemistry and Photobiology*, 88:215 – 222.

Siani, A.M., Casale, G.R., Diémoz, H., Agnesod, G., Kimlin, M.G., Lang, C.A., Colosimo, A. (2008) Personal UV exposure in high albedo alpine sites. *Atmospheric Chemistry and Physics*, 8:3749 – 3760

Siani, A.M., Casale, G.R., Sisto, R., Borra, M., Kimlin, M.G., Lang, C.A., Colosimo, A. (2009) Short-Term UV Exposure of Sunbathers at a Mediterranean Sea Site *Photochemistry and Photobiology*, 85:171 – 177.

Siani, A.M., Casale, G.R., Sisto, R., Colosimo, A., Lang. C.A., Kimlin M.G. (2011) Occupational exposures to solar ultraviolet radiation of vineyard workers in Tuscany (Italy) *Photochemistry and Photobiology*, 87:925 – 934.

Siani, A.M., Casale, G.R., Modesti, S., Parisi, A.V., Colosimo, A. (2014) Investigation on the capability of polysulphone for measuring biologically effective solar UV exposures. *Photochemical and Photobiological Sciences*, 13(3):521 – 530.

World Health Organization (2010). *Solar ultraviolet radiation. Assessing the environmental burden of disease at national and local levels*. WHO, Geneva, 17.

RECENT ADVANCES IN ASSESSING THE DYNAMICS OF PHYTOPLANKTON ASSEMBLAGES BY HIGH FREQUENCY ANALYSIS AT THE SINGLE CELL LEVEL

Denis Michel, Thyssen Melilotus, Dugenne Mathilde, Grégori Gérald

Université d'Aix-Marseille, CNRS/INSU, IRD, MIO, UMR 7294, 163 avenue de Luminy, Case 901, 13288 Marseille cedex 09, France
michel.denis@mio.osupytheas.fr, melilotus.thyssen@mio.osupytheas.fr, mathilde.dugenne@mio.osupytheas.fr, gerald.gregori@mio.osupytheas.fr,

Keywords : phytoplankton, dynamics of temporal and spatial distributions, in situ high frequency analysis, automated flow cytometry

Glossary : CNRS : Centre National de la Recherche Scientifique ; INSU : Institut National des Sciences de l'Univers ; IRD : Institut de Recherche pour le Développement ; MIO : Mediterranean Institute of Oceanography ; UMR : unité mixte de recherche ; CEL2SAT: acronyme of the project entitled "*Diversité structurelle et fonctionnelle du phytoplancton à l'échelle régionale de la Manche Occidentale : de la cellule à l'image satellite » ;* PRECYM : Plate-forme Régionale de Cytométrie pour la Microbiologie ; GIPREB : Groupement d'Intérêt Public pour la Réhabilitation de l'Etang de Berre.

Abstract
Phytoplankton plays a major role in oceanic biogeochemical processes by converting inorganic material into organic matter, the feeding source of the entire marine food chain. Its large diversity and short cell cycle make it very sensitive to environmental changes. To overcome difficulties linked to the very large size range and species number, and correctly account for phytoplankton dynamics in natural environments, it is critical to observe it at the single cell level, at an hour time scale, as well as at a spatial submesoscale. The availability of commercialised, dedicated automated flow cytometers (Cytosense), eventually submersible (Cytosub), enables to address the dynamics of phytoplankton temporal and spatial distributions.
The reported recent developments in that field stem from 9 years of experience and leadership. They concern investigations with a Cytosense conducted on a lagoon near Marseille, on a research vessel in the English Channel (coupled to a FerryBox), on a EOL (Environnement Observable Littoral, CNRS-Mobilis) buoy in the Bay of Villefranche sur Mer (France), in the Western Mediterranean Sea (2 cruises in 2013) and on the ferry Armorique connecting Roscoff (France) to Plymouth (UK) (coupled to a FerryBox). Except that in the English Channel, the Cytosense was remotely controlled. New developments under progress target heterotrophic microorganisms.

1 INTRODUCTION

Phytoplankton is the major producer of organic matter on earth, contributing up to 45% of annual global photosynthesis, though it only represents about 2% of the overall photosynthetic biomass (Field et al., 1998; Falkowski et al., 2003). Such a capacity stems from its existence essentially in the form of single cell organisms that can grow and divide rapidly (up to 2 divisions day^{-1} (Alpine and Cloern, 1988; Furnas, 1991) and spread over 70% of the earth surface. This short cell cycle makes phytoplankton able to react promptly (hour time scale) to any change in its environment. The current phytoplankton is the outcome of about 3 billion year evolution, from a time where life was only present in the form of single cell organisms. These organisms had to adapt to an infinite number of environmental changes, generating over years an extreme diversity that still remains to be fully unveiled. Because it is at the basis of the so called "food chain" in the ocean, being the major source of matter and energy for the higher trophic levels, in both pelagic and benthic environments (Li et al., 2006), phytoplankton was and still is the subject of intensive investigations. Many techniques were used. Optical microscopy is the only one giving access to cell identification, at least the larger (>20 µm) ones and is commonly used in phytoplankton monitoring. However, analysis by optical microscopy is time consuming, requires highly specialised staff and therefore cannot handle large numbers of samples. Consequently, it is unsuitable to investigate short-term variability or spatial distribution of phytoplankton. Pigment analysis by high-performance liquid chromatography (HPLC) opened the way to a better and faster quantification of

dominant groups (Jeffrey and Vest, 1997). The lack of pigment specificity makes interpretations difficult and limits the benefit of this approach even when coupled to optical microscopy (Havskum et al., 2004), particularly regarding the dynamics of phytoplankton. Fluorometry is also a bulk approach, and by dealing with optical signals, it is suitable for high frequency recording of the fluorescence that is emitted by chlorophyll *a* upon appropriate excitation. This is why fluorometry is easily and widely used to estimate the chlorophyll *a* content in water masses and to address its variability at short temporal and spatial scales (Rantarjavi et al., 1998). Serious limitations are linked to the conversion of chlorophyll *a* content to carbon biomass and also to the limited seawater volume generating the collected fluorescence signal. It means that large cells that are weakly concentrated may easily escape to observation.

Flow cytometry is a technique analysing at high rate optical properties of suspended individual particles flowing in a stream through an optical beam. When intercepted by the optical beam (usually a laser beam) each particle generates scatter and fluorescence (naturally or after staining) signals. Data analysis groups particles sharing similar optical properties and some instruments are able to sort out cell groups for further investigations. This technique was initially developed for the needs of the biomedical field. Phytoplankton being essentially represented by single cells in suspension in seawater and containing fluorescent photosynthetic pigments appeared suitable flow cytometry analysis. The potentialities of flow cytometry analysis and cell sorting were thus investigated in the early 1980s (Yentsch at al., 1983) and shortly after led to the breakthrough that represented the discovery of *Prochlorococcus* (Chisholm et al., 1988) that makes the largest group in the oligotrophic ocean (Partensky et al., 1999).

This finding fostered the development of new flow cytometers specifically designed for the analysis of phytoplankton in natural samples such as the optical plankton analyser (OPA ; Dubelaar et al., 1989 ; Jonker et al., 1995). This effort went on with the design of the Cytobuoy flow cytometer, conceived to be operated *in situ* in an automated way (Dubelaar et al., 1999). A specific feature of this instrument was the recording of the pulse shape of the signals generated by phytoplankton cells when they are intercepted by the laser beam, enabling the analysis of chains which is not possible with conventional flow cytometers. The number of cells composing a given chain is defined by the number of repeated shape-patterns representing the flow cytometric signature of each cell. Another automated flow cytometer was developed at Woods Hole Ocean Institute to investigate pico and phytoplankton (Olson et al., 2003) but did not include pulse shape recording.

The first commercialised submersible flow cytometer (Cytosub) developed by the Cytobuoy company (Cytobuoy.com) was bought by our group in 2004. This new approach was first validated by Thyssen et al. (2008a) then applied to investigate the dynamics of the phytoplankton assemblage in the Bay of Marseille (Thyssen et al., 2008b) and its spatial distribution in the north eastern Atlantic (Thyssen et al., 2009). The Cytosub was further involved in a mesocosm experiment in the frame of a Canadian programme (Thyssen et al., 2011).

Meanwhile, the Cytobuoy company kept upgrading its instruments (Cytosub, Cytosense) and included an image in flow device, particularly efficient for large (>20 µm) cells. The Cytosense, a lab-top flow cytometer, can be run in an automated way on pumped water which provides a large flexibility in coastal monitoring stations and in studies conducted on research vessels or ships of opportunity.

2 RECENT APPLICATIONS

2.1 Phytoplankton dynamics remotely monitored from a buoy

Our Cytosub was adapted by the constructor so that it could be run inside the EOL (Environnement Observable Littoral, CNRS-Mobilis) buoy deployed in the Bay of Villefranche sur Mer ((43.682°N, 7.319°E; north western Mediterranean Sea) by the Universe Science Observatory (OSU) of Villefranche sur Mer (Fig. 1).

Figure 1 View of the EOL (Environnement Observable Littoral, CNRS-Mobilis) buoy deployed in the Bay of Villefranche sur Mer (France). The water column depth at this location is about 100 m.

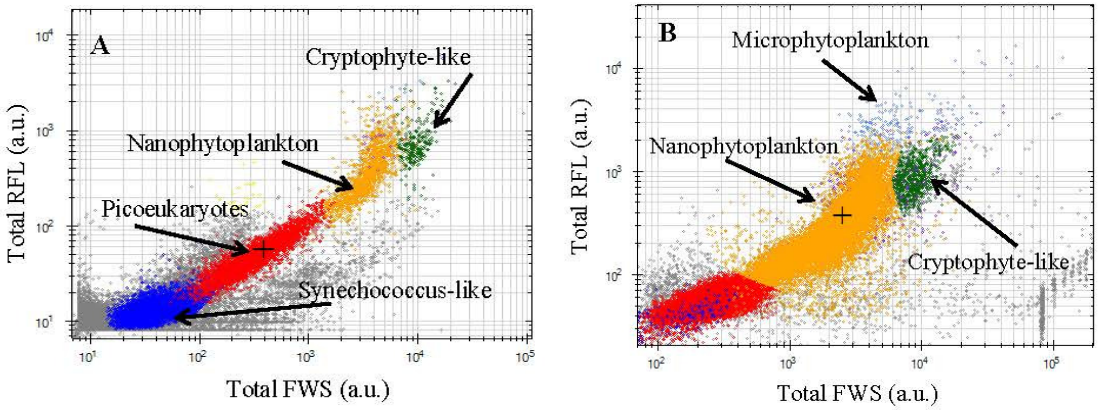


Figure 2 Typical cytogrammes illustrating the cluster resolution when using a low threshold (a) to focus on the smallest cells, and a higher threshold (b) to analyse a larger volume and account for bigger cells while skipping the smallest ones. Total FLR : integrated red fluorescence signal with respect to time. Total FWS: integrated forward scatter signal with respect to time. a.u. : arbitrary unit.

The energy was supplied by solar panels and stored on batteries located inside the EOL cabin, above sea level. The experiment lasted from January 24 to April 6 2012. The Cytosub analysed seawater pumped at 1 m depth. Thanks to a wifi connexion between the EOL buoy and the OSU, the Cytosub could be remotely controlled from our laboratory and data downloaded every day. Seawater samples were analysed every two hours by using in sequence two protocols, one focusing on the smallest cells, the other skipping the smallest cells to account for the larger ones by analysing more seawater as illustrated by Figure 2 displaying for each case cytogrammes of red fluorescence (emitted by chlorophyll a) versus forward scatter (linked to cell size). The key point was the trigger level set for the detection of the red fluorescence signal. A total of six phytoplanktonic cell groups were resolved throughout the validated 532 flow cytometry analyses.

The study, conducted in collaboration with scientists of the Laboratory of Oceanology of Villefranche sur Mer (LOV), succeeded in describing the onset of the spring bloom in the north western Mediterranean Sea and to document the influence of environmental pulse events on the *in situ* hourly-scale dynamics of the phytoplankton structure (Thyssen et al., 2014). By using a loess procedure, a controlled smoothing highlighted the global trend of the signals and of the abundance of each cluster. Three abundance pulses were observed and could be correlated to pulses of environmental variables (wind speed, rain fall, nutrient concentration). By working on

the periodic signals it was possible to characterise the cell cycle of the resolved clusters and differentiate their responses to the observed environmental changes (see Thyssen et al., 2014).

2.2 High frequency phytoplankton monitoring in surface waters

The first attempt to address the heterogeneity of phytoplankton surface distribution in the ocean with a Cytosub, clearly evidenced the dependence of the community structure on crossed water masses (Thyssen et al., 2009). In addition, it was possible to discriminate the periodic components from the longer-term trends in abundance so that cell cycles could be estimated. The large diversity of phytoplankton prevents to expect identifying the role of each species in biogeochemical processes. As a possible answer to this difficulty, the initial work on phytoplankton dynamics investigated at the single cell level revealed that cell clusters behaved as a single entity (Thyssen et al., 2008), and were recognised as functional response groups. This viewpoint was in line with the suggestion by LeQuéré (2005) to consider phytoplankton functionalities with respect to biogeochemical processes, each functionality being possibly shared by several species.
This approach was applied in particular to the dynamics of phytoplankton assemblages in the English Channel in the frame of the CEL2SAT (CNRS-INSU) project, by combining automated instruments, the Cytosense of the PRECYM regional flow cytometry platform of the MIO (http://precym.mio.univ-amu.fr) and a FerryBox instrument already installed on the ferry Armorique connecting Roscoff (France) to Plymouth (United Kingdom) twice a day. The coupling of the Cytosense with the FerryBox lasted from 16 May to 17 September 2013 and yielded technological and software developments that will be ready for use when implementing these instruments onboard ships of opportunity. The CEL2SAT project yielded an unprecedented amount of data on the phytoplankton distribution across the English Channel between Roscoff and Plymouth and the exploitation of the whole dataset is still under progress.

2.3 Response of the phytoplankton assemblage to a gust of wind.

The Berre lagoon, close to Marseille (France) is one of the largest brackish lagoons in the Mediterranean area with its 155 km^2 surface and makes a shallow water body (maximum depth: 9m). The construction of a hydroelectric power plant on its edge completely disturbed its ecosystem by frequent and large freshwater discharges that added to natural freshwater outflows. The lagoon is also under the influence of surrounding oil refineries and chemical factories. The seawater inflow goes through the Caronte channel and prevailing winds (northerly and easternly) make additional strong natural forcings. To alleviate the impact of the hydroelectric plant, its freshwater discharge was constrained with respect to time and volume. An agency (GIPREB), a public consortium for the restoration of the Berre lagoon, was set up to implement a detailed survey and monitor the lagoon eutrophication since 1995. However the short-term variation of the phytoplankton compartment was not taken into account in that survey. To fill this gap, we addressed the response of the phytoplankton to the large and sudden changes in its environment, particularly gusts of wind. To achieve this goal, we installed an automated flow cytometer (Cytosense instrument from the PRECYM regional flow cytometry platform of the MIO; http://precym.mio.univ-amu.fr) equipped with an image in flow device, inside the GIPREB building on the lagoon shore and ran hourly analyses of the lagoon water pumped 250 m off the shore, at 1 m above bottom. The experimental setup enabled parallel collection of hydrological variables on the same water. Subsamples from the same analysed water were taken at lower frequency to determine nutrient concentrations. Two flow cytometry protocols were run in sequence, one focusing on small cells, the other one on larger cells by setting an appropriate threshold on the red fluorescence signal (triggering the data acquisition) to skip the small cells and analyse a larger volume, bigger cells being less abundant. Cells sharing similar scatter and fluorescence properties were regrouped and thus we could define 12 distinct clusters labelled C1 to C12, C1 corresponding to the largest observed cells (56.4 ± 12.2 µm) and C12 to the smallest (0.9 ± 0.1 µm). Cells from C1 to C4 belonged to the microphytoplankton size class, those from C5 to C10 to the nanophytoplankton size class and cells of C11 and C12 were picoplankton. Most of the microphytoplankton cells were identified from their images. Figure 3 displays a typical cytogramme of red fluorescence (FLR) versus sideward scatter (SWS) in which clusters C4 and C5 are singled out with corresponding images of microphytoplankton cells. During the October 2011 experiment, a strong north wind (Mistral) event occurred with mean speed >20 ms^{-1} and the lagoon water temperature dropped by 5.2°C between 5 and 11 October. The meteorological data (wind speed and direction) were provided by the meteorological station of the Mediterranean Institute of Oceanography (MIO) located on the Frioul archipelago, in the Bay of Marseille. The high frequency observation of

phytoplankton cells enabled to characterise the dynamics of each resolved cluster before, during and after the wind event. One major result of this experiment was the confirmation of clusters as functional response groups as initially demonstrated by Thyssen et al. (2008b).

In this experiment, a particular attention was given to the impact of gusts of winds on specific growth rates. Because the cell optical properties are linked to cell size, high frequency single cell analysis provides a good alternative to estimate growth rates. Indeed, the periodic increase and decrease of the cell scatter-signal intensity expresses the cell growth and division occurring during the cell cycle (Durand, 1995 ; Binder et al., 1996). Sosik et al. (2003) developed a model to calculate growth rates of natural phytoplanktonic groups from the diel variations of the phytoplankton cell dimensions derived from automated high frequency flow cytometry surveys. This approach can be further improved by estimating cell dimensions derived from image in flow attachments (Sieracki et al., 1998). The Cytosense that we used in the Berre Lagoon experiment being equipped with an image in flow device, we took advantage of this capacity to estimate biovolumes. We adapted the size-structured matrix population model of Sosik et al. (2003) to process the single cell analysis of two dinoflagellate clusters and estimate their division rates before, during and after a gust of wind. Both clusters were composed of cells large enough to make images suitable for determining cell dimensions and abundant enough to yield robust results. Indeed, clusters with larger cells were made of cells in low abundance. Cells in one selected cluster were identified with *Gymnodinium sp*. from the collected pictures. Their mean cell biovolume increased one hour after dawn before it progressively decreased upon mitotic divisions. The amplitude of both processes varied before, during and after the wind event so that growth rates ranged between 0.22 d^{-1} (<1 division per day) and 0.85 d^{-1} (>1 division per day). In contrast, pictures could not help in identifying cells in the other cluster. Nevertheless, results showed that the ratio of max/min mean biovolume varied from 1.26 to 1.55 during the gust of wind, and growth rates progressively increased with regard to the variation of mean biovolume amplitude to reach 0.46 d^{-1} on 11 October. The striking result of this investigation was to demonstrate, thanks to the automated high frequency observation of phytoplankton, that cell physiological changes can prevail over the response of abundance. The full study is reported in Dugenne et al. (2014). As a major consequence, variations in phytoplankton growth rates would generate a mismatch between the related photosynthetic carbon fixation and net primary production derived from ocean color data. Indeed, usual low frequency (monthly or bimonthly) surveys cannot take into account pulse events and the subsequent phytoplankton biomass pulses that significantly contribute to biogeochemical fluxes and budgets at the annual scale (Lomas et al., 2009).

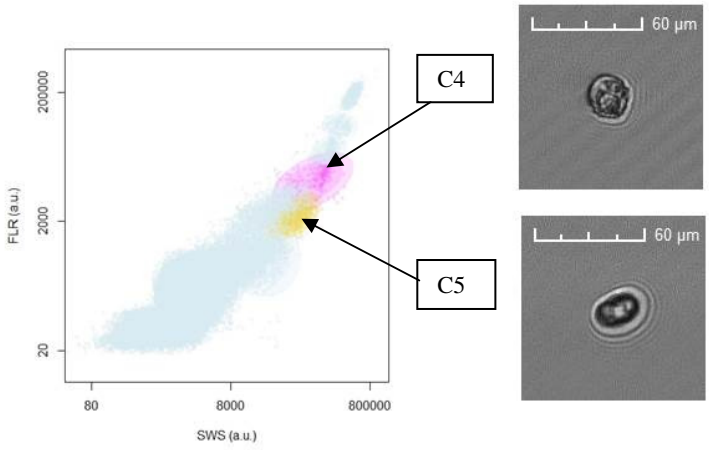


Figure 3 Typical cytogramme of the phytoplankton analysis in the Berre lagoon, displaying red fluorescence (chlorophyll) with respect to side scatter linked to cell structure. Pictures of cells composing clusters C4 and C5 are also shown. a.u. : arbitrary unit.

3 DISCUSSION

If the concept of automated flow cytometry analysis first emerged in the biomedical field (see Abu-Absi et al., 2003 and references herein) like flow cytometry itself, it was developed in parallel in the marine field to address the *in situ* dynamics of phytoplankton (Dubelaar et al., 1999; Olson et al., 2003), and further on in fresh water microbial ecology (Hammes et al., 2012; Besmer et al., 2014). Because we were the first to purchase the only commercialised automated and submersible flow cytometer (Cytosub, Cytobuoy bv) most of all the investigations reported above and conducted with the Cytosub and the new generation of Cytosense with image in flow made breakthroughs in the field of the phytoplankton dynamics. It started by giving ground to the concept of functional response groups for clusters resolved by flow cytometry. It went on by demonstrating the efficiency of this approach to conduct time series at high frequency (hour scale), to investigate spatial heterogeneity of phytoplankton at sub-meso-scale, to operate these instruments on pumped water, whether from a sailing boat, a ferry, a buoy, a research vessel. For the most recent experiments, they were run being remotely controlled. The great flexibility of this approach is now demonstrated and consequently open the way to new fields of investigation that were largely out of reach up to now. However, if the instrumentation gap can be considered as resolved now, one coming challenge is the automation of the data treatment. Indeed, the high frequency analysis is generating a huge amount of data and there is an urgent need for automated treatments. Some solutions were already explored (Malkassian et al., 2011) but more investigations are needed. The image in flow device re-establishes to some extent the identification capacity of optical microscopy, at least for large enough cells. This property was exploited to anticipate the development of harmful algae (Campbell et al., 2010). Opening the access to phytoplankton dynamics at the single cell level gives access to the functioning of this microbial community and its reaction to environmental changes. These automated flow cytometers are the only instruments able to generate biological information in an automated way and at a rate compatible with the dynamics of the observed microorganisms. It is expected that they will be soon implemented in the ocean observing systems to supply biogeochemical models with more robust biological information.

4 NEXT CHALLENGES

As presented in this report, automation of flow cytometry in the aquatic environment exclusively concerned phytoplankton because it is essentially composed of single cells in suspension in water and is able to generate fluorescence signals upon excitation of its photosynthetic pigments. In contrast, heterotrophic microorganisms in the aquatic environment require to be stained with fluorescent dies in order to be analysed by flow cytometry. This is currently achieved in the laboratory where this process implies more handling and more time with staining, incubation and analysis. To extend the automated high frequency flow cytometry analysis to aquatic heterotrophic microorganisms, it is necessary to automate the whole staining procedure. To solve this problem, in collaboration with the Cytobuoy company, we designed a new instrument (Cytopro) with an automated staining module, dedicated to heterotrophic prokaryotes and microzooplankton. The Cytopro is currently under evaluation and validation before being applied to the aquatic environment, both marine and fresh water. The coupling of Cytosense and Cytopro instruments will open the way to the investigation of the joint dynamics and interactions of autotrophic and heterotrophic microorganisms, another new field of research made accessible. The next big challenge ahead will be to make this instrument submersible to address the dynamics of heterotrophic microorganisms in the entire water column which is still out of reach but necessary to decipher this still largely unknown world (Aristegui et al., 2009). Meanwhile, other groups involved in microbial ecology in fresh water developed different solutions for the automation of the staining procedure and the flow cytometry analysis with conventional instruments, to address microbial dynamics in aquatic ecosystems (Besmer et al., 2014; Fontana et al., 2014). Finally, considering the needs generated by global ocean observing systems, a major challenge already mentioned in this report is the automation of the data treatment and the handling of the related huge data bases.
A really appealing and thriving new research domain is now made available to a number of scientists.

Note: this report is derived from a conference given by Michel Denis at the CISB (Centro Interdipartimentale per l'analisi dei modelli e dell'informazione nei sistemi biomedici) meeting "The CISB scientific activity: recent and seminal achievements". Rome, May 29-30, 2014, Palazzo Baleani. Melilotus Thyssen headed the CEL2SAT project.

ACKNOWLEGEMENTS

The reported research work benefited of several grants and supports over the past ten years:
CYMIS contract between CNRS and the City of Marseille, partially supported by "l'Agence de l'Eau Rhône Méditerranée Corse" ; MIO, Aix Marseille University and Provence Alpes Côte d'Azur (PACA) Council together with European Commission support (Fonds européen de développement régional (FEDER)), CNRS-INSU (MISE and CEL2SAT projects). M. Dugenne benefits from a PhD fellowship of the French Ministry of Higher Education and Research.

REFERENCES

Abu-Absi, N. R., Zamamiri, A., Kacmar, J., Balogh, S.J. and Srienc, F. (2003) Automated flow cytometry for acquisition of time-dependent population data. *CytometryA*. 51: 87–96. doi:10.1002/cyto.a.10016.

Alpine, A. E and Cloern, J. E. (1988) Phytoplankton growth rates in a light-limited environment, San Fransisco Bay. *Marine Ecology Progress Series* 44: 167-173.

Besmer, M. D., Weissbrodt, D. G., BradleyE. Kratochvil, B. E., Sigrist, J. A., Mathias S. Weyland, M. S. and Frederik Hammes, F. (2014) The feasibility of automated online flow cytometry for *in-situ* monitoring of microbial dynamics in aquatic ecosystems. *Frontiers in Microbiology, section Systems Microbiology*. doi: 10.3389/fmicb.2014.00265

Binder, B. J., Chisholm, S. W., Olson, R. J., Frankel, S. L. and Worden, A. Z. (1996) Dynamics of picophytoplankton, ultraphytoplankton and bacteria in the central equatorial Pacific. *Deep-Sea Research* II 43: 907-931.

Campbell, L., Olson, R. J., Sosik, H. M., Araham, A., Henrichs, D. W., Hyatt, C. J. and Buskey, E. J. (2010) First harmful *Dinophysis* (Dinophyceae, Dinophysiales) bloom in the U.S is revealed by automated imaging flow cytometry. Journal of Phycology 46: 66-75.

Chisholm, S. W., Olson, R. J., Zettler, E.R., Goericke, R., Waterbury, J. B. and Welschmeyer, N. A. (1988) A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* 334: 340-343.

Dubelaar, G. B. J., Groenewegen, A. C., Stokdijk, W., van den Engh G .J. and Visser, J. W. M. (1989) The optical plankton analyser (OPA): a flow cytometer for plankton analysis, II: Specifications. *Cytometry* 10: 529-539.

Dubelaar, G. B. J., Gerritzen, P., Beeker, A. E. R., Jonker, R. and Tangen, K. (1999) Design and first results of Cytobuoy: A wireless flow cytometer for in situ analysis of marine and fresh waters. *Cytometry* 37: 247–254.

Dugenne, M., Thyssen, M., Nerini, D., Mante, C., Poggiale, J.C., Garcia, N., Garcia, F., Pommeret, D. and Grégori, G. (2014) Consequence of a sudden wind event on the dynamics of a coastal phytoplankton community: an insight into specific population growth rates using a single cell high frequency approach. *Frontiers in Microbiology, section Systems Microbiology*. Doi: 3389/fmicb.2014.00485.

Durand, M. D. (1995) Phytoplankton growth and diel variations in beam attenuation through individual cell analysis. *PhD Thesis*, MIT/WHOI.

Falkowski, P. G., Laws, E. A., Barber, R. T. and Murray, J. W. (2003) Phytoplankton and their role in Primary, new, and export production. In M. J. R. Fasham.(ed.) *Ocean biogeochemistry*, Springer-Verlag Berlin Heidelberg, p 99-121.

Field, C. B., Behrenfeld, M. J., Randerson, J. T. and Falkowski, P. G. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281: 237–240.

Fontana. S., Jokela, J. and Pomati, F. (2014) Opportunities and challenges in deriving phytoplankton diversity measures from individual trait-based data obtained by scanning flow-cytometry. ***Frontiers in Microbiology, section Systems Microbiology***. Doi: 10.3389/fmicb.2014.00324.

Furnas, M. J. (1991) Net *in situ* growth rates of phytoplankton in an oligotrophic, tropical shelf ecosystem. *Limnology and Ocenaography 36*(1): 13-29.

Havskum, H., Schlüter, L., Scharek, R., Berdalet, E. and Jacquet, S. (2004) Routine quantification of phytoplankton groups-microscopy or pigment analyses? *Marine Ecology Progress Series,* 273: 31–42.

Jeffrey, S. W. and Vest, M. (1997) Introduction to marine phytoplankton and their pigment signatures. In S. W. Jeffrey, R. F. C. Mantoura, & S. W. Wright (Eds.), *Phytoplankton pigments in oceanography* (pp. 37–84). Paris: UNESCO.

Jonker, R. R., Meulemans, J. T., Dubelaar, G. B. J., Wilkins, M. F. and Ringelberg, J. (1995) Flow cytometry: A powerful tool in analysis of biomass distributions in phytoplankton. *Water Science and Technology* 32: 177-182.

LeQuéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., Christine Klaas, C., Kohfeld, K. E., Louis Legendre, L., Manizza, M., Trevor Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J. and Wolf-Gladrow, D. (2005), Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology* 11: 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.

Li, W. K. W., Harrison, W. G. and Head, E. J. H. (2006). Coherent Sign Switching in Multiyear Trends of Microbial Plankton. *Science* 1 1: 1157–1160. doi:10.1126/science.1122748.

Lomas, M. W., Roberts, N., Lipschultz, F., Krause, J. W., Nelson, D. M. and Bates, N. R. (2009) Biogeochemical responses to late-winter storms in the Sargasso Sea IV. Rapid succession of major phytoplankton groups. *Deep Sea Research I* 56 : 892-909.

Malkassian, A., Nerini, D., Van Dijk, M. A., Thyssen, M., Mante, C. and Grégori, G. (2011) Functional analysis and classification of phytoplankton based on data from an automated flow cytometer. *Cytometry A* 79A : 263-275.

Olson, R. J., Shalapyonok, A. and Sosik, H. M. (2003) An automated flow cytometer for analyzing pico and nanophytoplankton = FlowCytobot. *Deep Sea Research Part I* 50: 301–315.

Partensky, F., Blanchot, J. and Vaulot, D. (1999) Differential distribution of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. In: Charpy L, Larkum A.W.D. (eds), Marine cyanobacteria. *Bulletin de l'Institut Océanographique de Monaco* No sp. 19: 457-475.

Rantajarvi, E., Olsonen, R., Hallfors, S., Leppanen, J.-M. and Raateoja, M. (1998) Effect of sampling frequency on detection of natural variability in phytoplankton: Unattended high-frequency measurements on board ferries in the Baltic Sea. *ICES Journal of Marine Scienc,*55: 697–704.

Sieracki, C. K., Sieracki, M. E. and Yentsch, C. S. (1998) An imaging-in-flow system for automated analysis of marine microplankton. *Marine Ecology Progress Series* 168 : 285-296.

Sosik, H. M., Olson, R. J., Neubert, M. G. and Shalapyonok, A. (2003) Growth rates of coastal phytoplankton from time-series measurements with a submersible flow cytometer. *Limnology and Oceanography* 48: 1756-1765.

Thyssen, M., Tarran, G. A, Zubkov, M. V., Holland, R. J., Grégori, G., Burkill, P. H. and Denis, M. (2008a) The emergence of automated high frequency flow cytometry: revealing temporal and spatial phytoplankton variability. *Journal of Plankton Research* 30: 333-343.

Thyssen M., Mathieu D., Garcia N. and Denis M. (2008b) Short-term variation of phytoplankton assemblages in Mediterranean coastal waters recorded with an automated submerged flow cytometer. *Journal of Plankton Research* 30: 1027-1040.

Thyssen M., Garcia N. and Denis M. (2009) Sub meso scale phytoplankton distribution in the North East Atlantic surface waters determined with an automated flow cytometer. *Biogeosciences* 6: 569-583.

Thyssen, M., Ferreyra, G., Moreau, S., Schloss, I., Denis, M. and Demers, S. (2011) Combined effects of ultraviolet radiations B and temperature increase on phytoplankton dynamics and cell cycle using pulse shape recording flow cytometry. *Journal of Experimental Marine Biology and Ecology* 406: 95-107.

Thyssen, M., Grégori, G. J., Grisoni, J.-M., Pedrotti, M.-L., Mousseau, L., Artigas, L. F., Marro, S., Garcia, N., Passafiume, O. and Denis, M.J. (2014) Onset of the spring bloom in the northwestern Mediterranean Sea:

influence of environmental pulse events on the in situ hourly-scale dynamics of the phytoplankton community structure. *Frontiers in Microbiology, section Systems Microbiology*. doi: 10.3389/fmicb.2014.00387

Yentsch, C. M., Horan, P. K., Muirhead, K., Dortch, Q., Haugen, E., Legendre, L., Murphy, L., Perry, M. J., Phinney, D. A., Pomponi, S. A., Spinrad, R. W., Wood, M., Yentsch, C. S. and Zahuranec, B. J. (1983) Flow cytometry and cell sorting: A technique for analysis and cell sorting of aquatic particles. *Limnology and Oceanography* 28: 1275-1280.

Modeling the Nervous System by Evolutionary Connectionism

Neri Accornero and Marco Capozza
Neuroscience Department, Sapienza University, Viale dell'Università 30, 00185, Rome, Italy
neri.accornero@gmail.com, marco.capozza@gmail.com

Keywords: Artificial Neural Network, Connectionism, Simulation, Evolution, Genetic Algorithms, Nervous System

Abstract: Artificial neural networks (ANN) are excellent biological-like models to simulate nervous system functions. Their spontaneous biological likeness can be further enhanced by letting biological-like procedures, such as evolutionary selection, shape the investigated ANN models. This makes the investigator duties shift from designing the networks to designing the environment where the networks behave, so reducing the impact by possible investigator preconceptions on the models. In this paper we present a few examples about modeling nervous system functions with ANN subjected to evolutionary selection by genetic algorithms. This methodology, called Evolutionary Connectionism, proved feasible and profitable in studying both sensorimotor and cognitive nervous functions through simulation.

1 INTRODUCTION

Artificial neural networks (ANN) are the best model available so far to simulate and understand biological nervous systems, inasmuch as they spontaneously and effectively tend to reproduce essential nervous functions and features, even if in a necessarily simplified way. In this paper we present a summary of some ANN models we developed, to show how powerful this methodology is in simulating nervous system functions, from simple motor reflex up to cognitive behaviour (see Accornero and Capozza, 2009 for a more thorough description on these and other models).

ANN intrinsic biological plausibility can be further increased by letting the modelled networks get shaped by genetic algorithms (Goldberg, 1989; Schaffer, Whitley and Eshelman, 1992) selecting through Darwinian evolution ANN populations adapted to perform required tasks ('genotypic learning'). This makes the investigator duties shift from designing the networks to designing the environment where the networks behave, so reducing the impact of possible investigator preconceptions on the model. This methodology, called Genetic Connectionism (Chalmers, 1990) or Evolutionary Connectionism (Calabretta and Parisi, 2005) and its models called Evolutionary Artificial Neural Networks (Yao, 1993, 1995, 1999), proved to be powerful and effective as well as compatible with actual biological reality.

The first two models we present here with this methodology produce, in an automatic and unsupervised way, ANN populations adapted to perform tasks needed to 'live' in their artificial environment. The first model (ALIFE) simulates a simple ecosystem where agents develop a motor strategy to achieve energy sources (food) in order to live and reproduce. In the second model (GAZE), an ANN population evolves that exhibits a human-like head-eye motor control to pursuit visual targets.

The third model (ARM) does not make use of genotypic learning, inasmuch as its purpose is to introduce a special, predesigned sensorimotor ANN model, which we called the 'triplet-net' model and we consider pivotal to model intentional movement and even operative thinking. This simulation concerns intentional movement, and it supports the 'ideomotor principle' (IMP) (James 1890, Kiesel and Hoffmann 2004, Stock and Stock 2004), i.e. the notion positing that voluntary actions are cognitively represented by their sensory effects and produced by anticipation (desire) of those effects.

In the fourth model (ACON) we extend the triplet-net model from the sensorimotor domain to the cognitive domain, to show once again by evolutionary connectionism how imagination spontaneously evolves as a way to reach goals. This can be interpreted as a beginning of 'thinking' in an artificial system, and it provides us

with a better understanding for mental functions critical to consciousness, such as recalling memories, foreseeing and reasoning.

2 'ALIFE' (ARTIFICIAL LIFE) SIMULATION

2.1 Introduction

This was our first simulation with evolutionary connectionism methods, dating back to 1994, DOS and VGA age. With it we aimed to preliminarily test the efficacy of genetic algorithms applied to ANN populations, before venturing to more demanding experiments.

2.2 Methods

This simulation consisted of an environment, 'world', with 'food' packets and a population of 'agents'. These were small neural networks, subjected to Darwinian evolution, which survived and reproduced by developing sensory and motor skills useful to achieve food.

The simulation world was a two-dimensional grid with 640x480 cells, a folded toroid. In this world agents consumed energy to stay alive and move, and each of them could replenish its energy by 'eating' food. There was no limit to the number of agents that could coexist in a same cell.

Time was discretized and measured in turns. Agents acted once in each turn, where 'acting' meant: receiving visual information, activating the neural network units, moving, eating food i.e. gaining energy, and possibly interbreeding, reproducing, or dying. Subsequent agent generations might coexist and overlap.

Food packets reflected 'light' diffused in the environment, so that agents with visual ability might perceive them, and agents with motor ability could reach them. All food packets contained a same energy amount. A food packet got 'eaten' when an agent came close to it within a given distance. A number of food packets, fixed or cyclically variable in time (seasonal cycles) at investigator's choice, was released into the environment each turn. Every new packet was introduced in a random cell within a predetermined maximum distance from a random other packet already present. Since both maximum horizontal and vertical distance were decidable by the investigator and might differ, the investigator was able to affect the shape and orientation for food clusters (disk clusters, vertical or horizontal bands, etc.). The total amount of food packets in the environment changed each turn, being affected by agent skill to reach them and seasonal pattern.

Agents (max. 200) were constituted by a variable number of neural units (cells) ranging from 1 to 12, arranged as an annulus ('body'), so that all agents had a circular shape with the same size but a possibly different cell number. Any cell might evolve visual ability, motor ability, both, or none; cells with neither visual nor motor ability actually acted as hidden neural units (Rumelhart, Hinton and McClelland, 1986; McClelland and Rumelhart, 1986). Cells mutually connected through connections. Cell abilities and connection weights were selected by evolution.

Sensory cells were activated by light reflected from food, with activation values depending on light intensity (decreasing with distance) and cell visual sensitivity. The area 'seen' by each cell was a circle sector whose orientation depended on the cell position within the agent body. This visual system had no 'retina', just simple directional light sensors, although several sensors could be placed side by side like insect ommatidia.

Each motor cell generated a force vector tangent to the agent body contour with strength depending on cell motor ability and activation. Rotation and translation resulted from the overall composition of all motor vectors on the agent body. Since one single vector could only make an agent rotate, in order to translate agents must activate at least two motor cells at the same time in a coordinated manner. Agent energy consumption to move depended on its total number of motor cells, their motor ability, and their activation levels. Agents consumed energy not only to move but also to stay alive (metabolism).

All agents were initially provided with the same energy content. If an agent energy dropped to zero due to metabolism and movement consumption with insufficient food intake, that agent died and was removed from the environment. If, on the contrary, food eating increased the agent energy over a given threshold, then the agent was reproduced and a mutated copy of it was released into the environment. There was no fitness score other than agent energy levels, which affected 'real-time' (simulation time, of course) agent lifespan and replication. All features constituting an agent (its number of cells, cell position in its body, cell visual and motor ability, and cell connection weight) were coded as a 0-and-1 string in the computer memory. This string represented the agent 'genome'. When an agent reproduced, a new copy of its genome was created with some

bit mutated, and a new agent was generated from the new genome. When two agents became closer within a given distance they interbred, meaning that they exchanged genome string segments ('genetic crossover'). All consequences of this genome exchange were immediately applied to their body structures, similarly to DNA exchanges occurring in bacteria.

The simulation computer program allowed the investigator to determine online all parameters involved in what we described above. The program also allowed viewing and analysing population performance through three different displays: a 'history' display (Fig. 1 left side), reporting in color code the number of agents along time, their overall ability to achieve food, their prevalent motor strategies (rotation and translation amplitude), and food availability; a 'world' display, where the entire world grid was displayed real-time, with agents shown as moving colored dots (color depending on the agent energy reserve), and food packets as small motionless white dots; and an 'individual' display (Fig. 1 right side), where a single agent selected by the investigator was shown real-time along with its body, cells, energy content, visual perceptions and motor actions.

We conducted several simulations using different parameter values. Simulations ended at the investigator's choice, usually after 6-7000 turns.

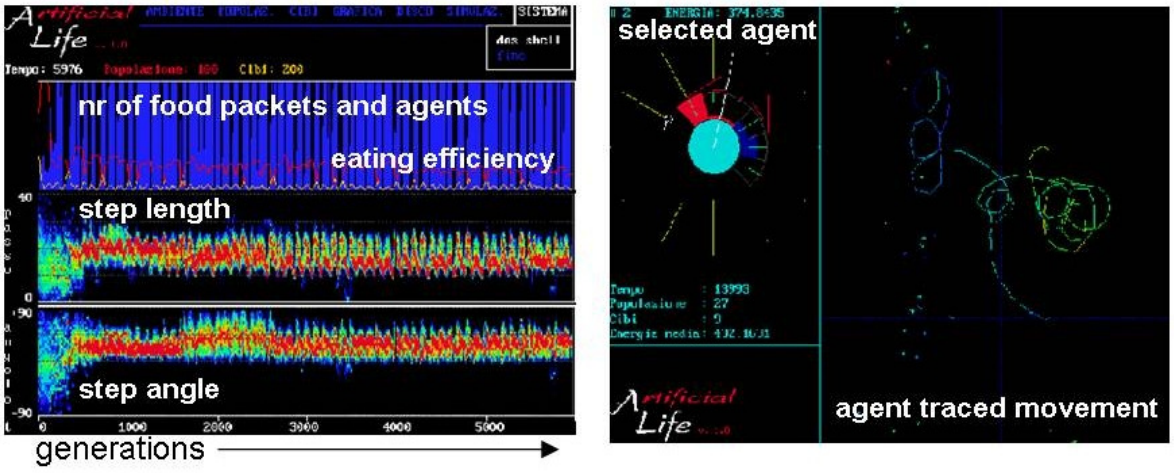


Figure 1: Historical (left) and individual (right) display offered by the ALIFE simulation program.

2.3 Results

Our aim to find the emergence of a working 'survival behavior' in the simulated populations was achieved almost invariably, even with different 'world' parameters. In all simulations an agent population with sensorimotor reflexes invariably evolved. Agent bodies were composed by sensory cells and motor cells, often with two or three sensory cells grouped at a body pole (which can be regarded as the rostral pole), and two, three, or four motor cells usually located on the left and right body sides in order to obtain a forward movement and the ability to steer.

2.4 Conclusions

This preliminary study showed that even simple genetic algorithms were highly effective in producing ANN adapted to a designed environment, similarly to biological evolution for real nervous systems. It also highlighted that the similarity extent between artificial ecosystems and the real world heavily depended on realism in simulating real physics laws.

3 'GAZE' SIMULATION

3.1 Introduction

This study aimed to verify whether genetic algorithms could assemble ANN able to control visual reflex movements in a simulated head-eye model. This head-eye model, even if simplified, was simulated more realistically than agent bodies in the ALIFE experiment, while the genetic algorithm was more abstract and formal than the simple 'eat food or die' principle used in ALIFE.

3.2 Methods

This simulation concerned a simplified two-dimensional head-eye model driven by an ANN to keep the eyes gazing at a moving target. The head-eye system (Fig. 2) consisted of a head rotating on the horizontal plane due to actions by two antagonist muscles simulating the left and right sternocleidomastoid muscles, and two eyes also rotating on the horizontal plane due to actions by two pairs of antagonist muscles (one pair each eye) simulating the medial and lateral eye muscles. Muscles were simulated by Hill spring model (Shadmehr and Arbib, 1992), with their relative length and insertion points acceptably approximating the real ones.

A horizontal bar-shaped target randomly moved with different velocities on a 430 x 888 pixel horizontal plane in front of the eyes, and it sometimes also emitted a virtual sound. The eyes perceived the target through two linear retinas, each one with a fovea. Target sounds were perceived by two ears placed on the left and right head side.

The ANN (Fig. 3) received visual input from the retinas, acoustic input from the ears, proprioceptive inputs on head and eye directions, and proprioceptive input on muscle tensions, for a total of 29 input units. From these inputs the ANN generated output commands for the head and eye muscles (6 output units). Besides input and output units, the ANN also included a varying number of hidden units ranging from 0 to 64, that received connections from the input units and other hidden units, and sent connections to the output units and other hidden units. The input units were simple linear units; all other units were classic sigmoid units (Rumelhart, Hinton and McClelland, 1986), having an analogic output ranging from 0 to 1 and equipped with modifiable learning bias. All features of any ANN were coded as a 0-and-1 string representing that ANN 'genome'.

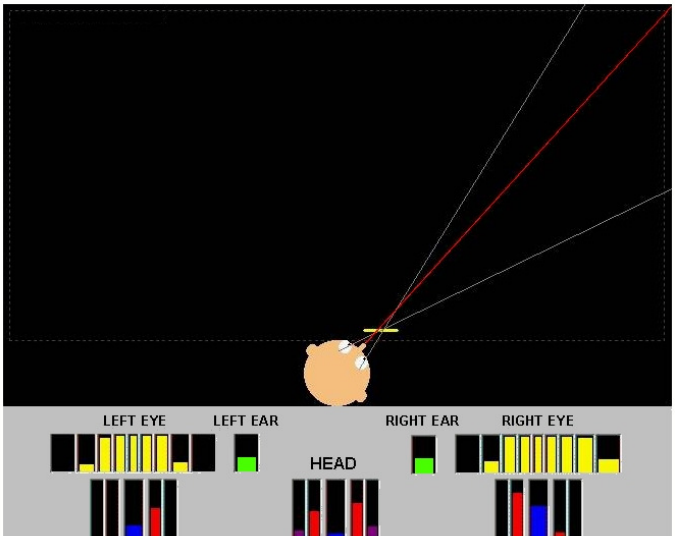


FIGURE 2: HEAD-EYE SYSTEM FOR THE GAZE MODEL. THE COLOURED BARS SHOW ACTIVATION LEVELS OF AGENT SENSORS AND EFFECTORS. YELLOW: VISUAL RETINA INPUT; GREEN: ACOUSTIC INPUT; BLUE: HEAD AND EYE DIRECTION PROPRIOCEPTION; PURPLE: MUSCLE TENSION

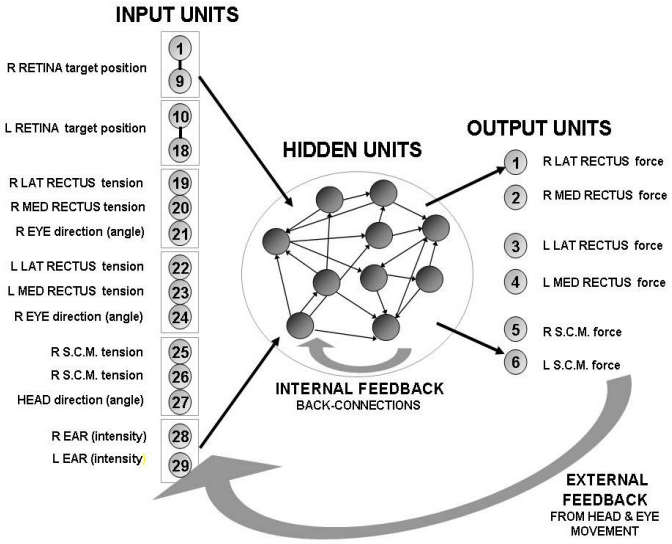


Figure 3: Artificial neural network for the GAZE model. Not all connections are depicted; actually, all input units send connections to all hidden units, and all hidden units to all other hidden units and to output units. R = right; L = left; LAT = lateral; MED = medial; S.C.M. = sternocleidomastoid muscle.

Simulation time was discretized and measured in turns. In each turn the target moved once and an agent acted once, where 'acting' meant: receiving visual information, activating the neural network units, and moving the head and eyes.

We set up 100 agents slightly different in their genomes and hence in their neural networks. When the simulation began, each agent was provided with a random number of hidden units ranging from 8 to 32, and the connection weights and the unit biases were initialized with random values. Then every agent was tested about its ability to keep gazing at the moving target during 10000 consecutive turns representing the agent lifespan. Each agent was tested individually; agents did not interact as they did in ALIFE. Three error values were measured in each turn as the angular distance between the direction of the target with respect to the agent and the direction where the agent head and eyes pointed to. Those values were recorded and added into three running (mis)fitness scores (one for the head and two for the eyes) for that agent. Once all 100 agents had been tested for 10000 turns each, a generation was ended, and the agents were compared each other on their three fitness scores singularly and ranked on their overall number of winning results (a non-parametric fitness evaluation). Agent genomes were then either reproduced into the next generation or discharged according with their rank (two offspring from the best agent, no offspring from the worst). Point mutation and genetic crossover were randomly applied to the offspring, with bit mutation rate = 0.0001 and crossover rate = 0.5.

In this simulation subsequent generations did not overlap. The length of the bar-shaped moving target randomly changed across generations. Target shape and trajectory were the same for all agents in a generation. Simulation ended when the mean angular error of the best agent in a generation dropped below 2 degrees for both the head and the eyes. At this point the best agent underwent specific tests with the target moving along pre-defined, non-random trajectories with fast and slow movements. These additional tests were also repeated with the agent head locked so that the agent could only move its eyes to keep gazing at the moving target.

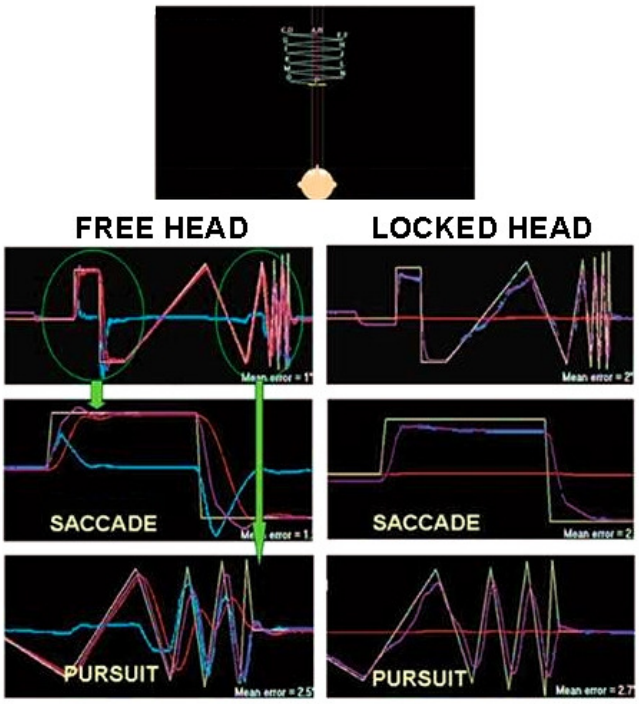


Figure 4: Traces of GAZE fast (saccades) and slow (pursuit) movements following fast and slow target displacements, with the head free to rotate (left column) or locked (right column). Yellow trace = target; red = head; blue = eyes; purple = gaze.

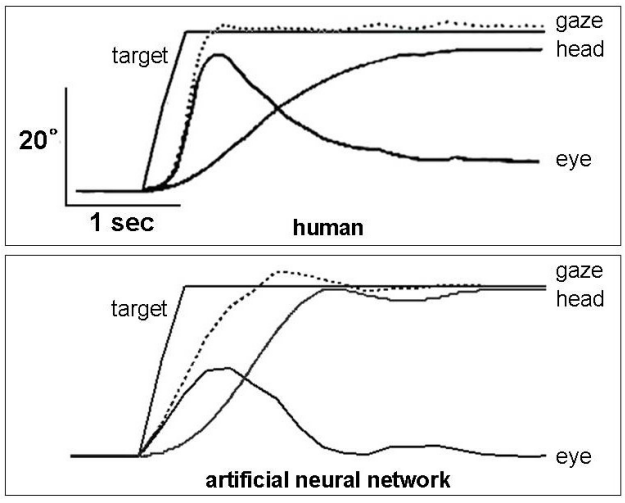


Figure 5: Comparison between human and GAZE movement patterns following a fast target displacement.

3.3 Results

The designed system proved able to develop the desired motor control. In about 4000 generations the agents became able to keep their head and eyes gazing at the target with a mean error less than 5 degrees. Specific tests (Fig. 4) showed that the system could generate rapid conjugated (binocular foveation) and unconjugated (vergence) movements, and fast

(saccades) and slow (smooth pursuit) movements following target position and velocity. When the same tests were repeated with the agent head locked (Fig.4 right), the system compensated efficaciously with ample eye movements in order to track the target, just like biological systems do.

We also observed (Fig. 5) that, after fast target movements, the eyes moved so that gaze rapidly approximated the new target position; shortly after, the head began to move, while the eyes made a compensatory movement backward into primary position, so that the gaze stayed fixed onto the target. This movement pattern closely resembled those exhibited by biological nervous systems in humans and monkeys (Bizzi, 1974). On the contrary, when the target movement was slow enough so that the head could follow, no eye movements were observed, which was similar once again to what happens in biological systems.

3.4 Conclusions

This study confirmed the effectiveness of genetic algorithms applied to ANN, and showed that applying realism at an even limited extent in simulating the biophysical reality of actual biological systems may result in the appearance of emergent behaviors closely resembling the real ones. Our results from the locked head test also demonstrated that these emergent properties even include the ability to produce appropriate behaviors in conditions never previously experimented by the population in evolution. This ability to generalize on limited experience is a striking feature of ANN models and biological nervous systems, and it seems even more likely to appear in ANN simulations with ANN models shaped by biological-like processes such as evolutionary selection.

4 'ARM' SIMULATION

4.1 Introduction

The ideomotor principle (IMP) (James 1890, Kiesel and Hoffmann 2004, Stock and Stock 2004) claims that the nervous system initiates voluntary actions by anticipating their typical sensory consequences. In this study we present an unsupervised ANN system that is as simple and basic as possible and learns to move a three-joint arm in a two-dimensional workspace using this principle. We examine its main features and compare them with those of human motor learning.

4.2 Methods

4.2.1 Model Design

This model does not make use of evolutionary connectionism and genotypic learning. Instead it makes use of phenotypic learning, because its purpose is to introduce a special, *predesigned* sensorimotor ANN model, which we called the 'triplet-net' model and we consider pivotal to model motor and behavior learning, intentional movement, and even operative thinking. Therefore in this simulation ANN are not arranged in an evolving population, and there are no subsequent ANN offsprings, but there is just one single ANN instead learning to move a limb in an unsupervised fashion on its own motor experience. This 'triplet-net' model will be then brought back to genotypic learning and evolutionary connectionism in the next ('ACON') model.

Our simulation consisted of an ANN controlling a three-joint simulated limb moving in a two-dimensional plane (Fig. 6). The network received sensory information on the limb position on its input units, and sent limb commands from its output units. Three vectors defined each limb movement: the initial sensory state S1 (before the movement); the final sensory state S2 (after the movement); and the neuromuscular activations M needed to pass from the initial to the final state. The S1 and S2 vectors were given to the ANN input units, and the ANN had to compute the M vector on its output units. Each S1 and S2 vector element activated one ANN input unit, and each ANN output unit generated one M element, so that each S1 and S2 element had a one-to-one mapping with the activation of one input unit, and each M element with the output of one output unit. Therefore we will speak of S1, S2 and M elements also as S1, S2 and M units.

Because IMP states that intentional limb movements depend on anticipation of their sensory effects, the ANN input units receiving after each movement sensory information on the *final* limb state (S2 units) also received, before each movement, motor commands from a component outside the network that established

where the moving hand should be positioned and therefore acted as 'Motor Will'. Commands from it to the ANN consisted of sensory representations of the desired final hand position, coded as visuospatial coordinates in agreement with the observation that motion planning in human takes place in the visually perceived space (Flanagan and Rao, 1995; Shadmehr, 2005). Unlike S2 units, the ANN input units receiving sensory information on the *initial* limb state (S1 units) did not receive motor commands, they only received 'proprioceptive' sensory information from the limb joint angles.

Given that velocity information was not indispensable to the key IMP mechanism as long as the limb was assumed to start from still and end still, we decided to give the ANN only sensory information about limb position (joint angles and spatial hand position).

4.2.2 Limb

The limb was designed to represent a simplified model of the human right arm comprising three segments, 'arm', 'forearm' and 'hand' articulated with three joints 'shoulder', 'elbow' and 'wrist', with the shoulder situated in a fixed point in space, and the hand able to move freely in the reachable space

4.2.3 Neural Network

The ANN was a two-layer neural network comprising 5 input units and 6 output units, fully connected with anterograde connections from input to output. There were no hidden units. The first three input units (S1 units) received 'proprioceptive' information on the opening angle for each of the three joints, normalized between -1 and 1. Before any movements these units received their information from the sensory pathways, and they kept that information alive up to the learning phase occurring after the movement (see below, 4.2.6). The last two input units (S2 units) received from the sensory pathways 'visuospatial' information on the hand position, encoded in polar coordinates. Immediately before a movement (movement phase, 4.2.5 step 2) this hand visuospatial information on S2 units was overwritten by Motor Will with activations corresponding to a new desired hand position (a motor command).

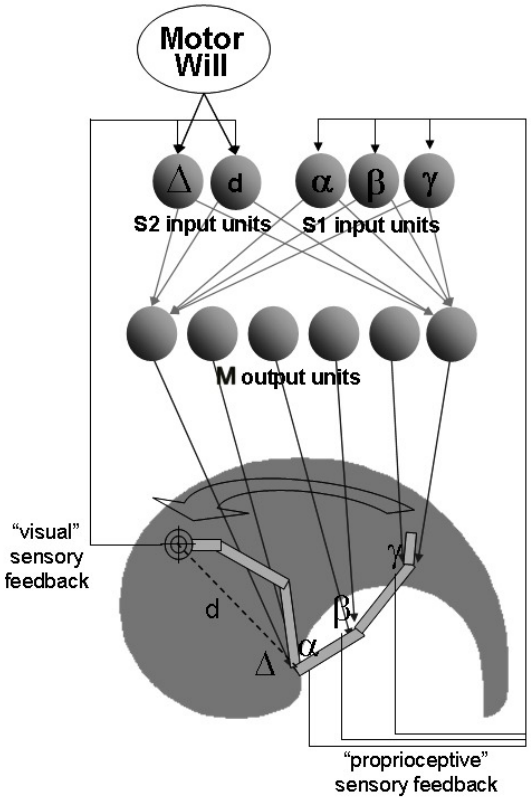


Figure 6: General architecture for the ARM model. The artificial neural network (ANN) controls a 3-joint limb moving in a two-dimensional plane. The ANN receives sensory feedback information on the limb and motor commands from a Motor Will.

4.2.4 Simulation Flow

When the simulation began, the connection weights and the output unit biases were initialized with random values ranging from -0.25 to +0.25. The arm was positioned with all the joints partly opened. After the initialization stage, the simulation proceeded in turns, each turn comprising the two phases, movement and learning, each comprising three steps.

4.2.5 Movement

1. The input units received sensory information from the arm: S1 units received the angles from the three joints, and S2 units the actual spatial location of the hand.
2. Motor Will overwrote S2 input unit activations with activations corresponding to a random desired hand position.
3. The input units activated the output units, and the joint opening angles therefore changed. The actual output values were recorded for use in the ensuing learning phase, during which they yielded the desired output, *target*

activations. The spatial error in pixels between the desired (target) and actual hand position is recorded to evaluate the ANN performance and not to assess motor learning.

4.2.6 Learning

1. The sensory pathways conveyed to the S2 units information on the new hand position.
2. The input units activated the output units again, this time using the new activation values on the S2 units corresponding to the hand position actually reached. These new outputs did not cause arm movements, they served only for learning. These new outputs were the ones the network would produce if the desired movement were actually the movement achieved in movement phase 4.2.5 step 3. The difference between the current outputs and the outputs recorded in that phase was the error to minimize during learning.
3. A standard delta rule (Rumelhart, Hinton and Williams, 1986) was applied to minimize the error vector calculated in the former step.

It should be noted that after the movement only S2 unit activations were updated with the new hand position, while S1 unit activations maintained the values corresponding to the joint angles before the movement. This was essential to the model functioning: at the beginning of the learning phase, S1 units must code the initial state before the movement, S2 units the final state after the movement, and M units the neuromuscular activations that caused the transition from the initial state to the final state. It was not even essential that S1 and S2 units encoded different sensory information (in our simulation, joint angles and hand position respectively): actually, *both* S1 and S2 units might encode *both* the joint angles and hand position. We made S1 and S2 units encode different information just to make the ANN task not too trivial. What was essential to the model was not the type of sensory information given to the S1 and S2 units, but that the learning phase would begin with the S1 units encoding *some* information on the arm state *before* the movement, the S2 units *after* the movement, and the M units the neuromuscular activations causing the movement. This was a triplet taken on experience on the real world, and it was what the ANN had to learn, so that whenever that same S2 final position would occur as a desired position on a same S1 initial state, the ANN would be able to generate those M activations bringing the arm from the S1 state to the S2 state.

4.2.7 Tests

Besides evaluating the 'online' spatial error after every movement (section 4.2.5 step 3), after every 5000 movements the program submitted the network to an 'offline' test entailing a predefined set of 588 target points (Fig. 7A) commanded by the Motor Will. During testing, the learning phase (section 4.2.6) was skipped. For each of the 588 points the position actually reached by the hand and the corresponding spatial error were recorded for later evaluation offline.

We conducted several simulations. In some simulations we introduced a sort of 'sensory blind spot', a wide circular area, covering up to 50% of the workspace and differentially positioned in the various simulations (Fig. 8A), where we skipped the learning phase when the hand ended up in this area.

In other trials, to assess whether learning depended on precise physical values inherent to the system, and to verify whether the controller system adapted to changes in the controlled system, we varied the sensory code used for hand position or the limb segment mass, right from the beginning or after advanced learning (30000 movements).

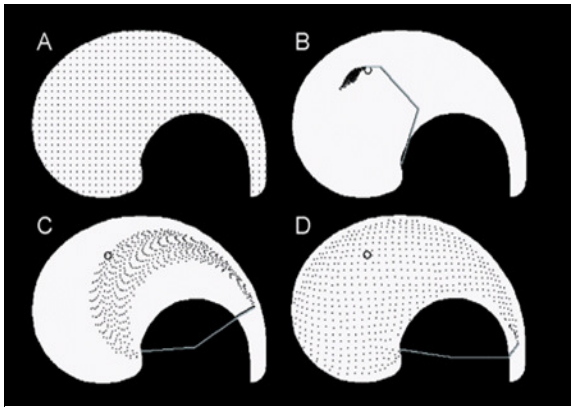


Figure 7: Progressive improvement in performance during the 588 test movements with motor experience. Small circle = hand starting point; black points = hand movement arrival point. A: target points; B: points effectively reached before learning; C: after 5000 random movements and D: after 30000 random movements.

4.3 Results

In all the simulations the tested ANN system improved from a mean spatial error of more than 150 pixels when simulation began to an error of less than 15 pixels after 10,000 movements (few tens of seconds on a modern pc) and fewer than 7 pixels after 30000 movements.

After 30000 movements, spatial error distributions showed that the system performed well over the whole workspace, except in the extreme tail in the drop-shaped area corresponding to extreme extension (Fig. 8B). The sensory blind spot had scarce influence on learning improvements (Fig. 8C, D). These results remained uninfluenced by the hand sensory code used, nor did they significantly suffer from mass changes in limb segments, before or after motor learning.

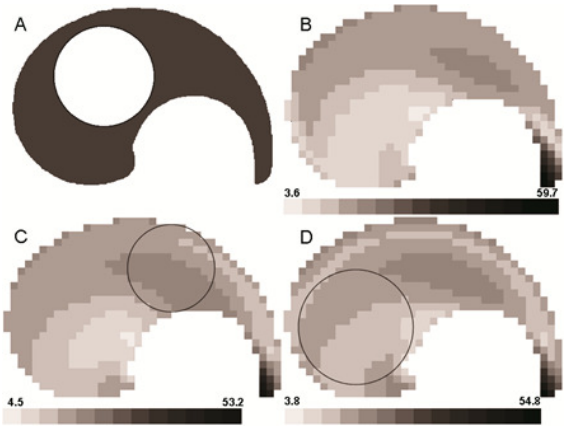


Figure 8: Spatial error distribution for the 588 test movements after 30000 random movements with and without 'sensory blind spot'. A: workspace area (dark grey area) with a generic blind spot (white disk); B: errors (in grey color code) without the sensory blind spot; C, D: with the sensory blind spot (black outline circle) in two different sizes and positions. Values are for spatial error in pixels.

4.4 Discussion

The simplified ANN simulation, focusing on the basic IMP features insofar as motor commands and sensory feedback reach the same S2 input units, effectively learned to move the arm in the workspace. It learned acceptably well even when we varied influential experimental variables such as the sensory code used for hand position, the mass for the limb segments to move, and when the ANN was able or unable to receive sensory feedback about movements performed in the workspace (sensory blind spot).

Our IMP model reproduced with acceptable approximation some human motor learning properties, such as learning from experience, ability to work regardless of the specific body segment features, ability to adapt to changes in these features, and the fact that even randomly-generated movements contributed to learning (infantile motor babbling). Like the human motor learning system, our ANN underwent completely unsupervised learning. We never used external sample sets, the ANN itself generated learning examples from its random movements and errors. The difference between output unit activations in two different functional phases (movement and learning) was utilized as the learning error, and those values were completely and locally available to the net.

In this model *the movement learned is not the desired movement but the movement effectively done*. The system nevertheless succeeded in performing with reasonable precision even movements never previously done and those finishing in the sensory blind spot. This system ability evidently stems from an ANN's well-known ability to generalize (Caudill and Butler, 1992), a feature allowing our ANN to interpolate and extrapolate information from the movements done, thus filling in unexplored movements and forming the general sensorimotor map valid for all movements.

In this simplified model the ANN consists of three unit groups coding the initial limb sensory state, the final state, and the action causing the limb to pass from its initial to its final state, that we will henceforth call a *sensorimotor triplet*, or simply *triplet*. This triplet model can be extended from elementary movements to more complex behaviors thus unifying the various intentional movement scales under a single principle. It involves several triplet-networks, linked so that the output units for each preceding triplet-net also act as the S2 input units for the ensuing triplet-net ('chained triplet-net' model). The S1 and S2 input units can receive sensory information not only from the musculoskeletal system, but from the whole body and external environment. In this chain, the S2 input units on the first net receive the actions desired by Will (actions that are more abstract than the simple and concrete desire to bring the hand to a desired position), and the ensuing nets progressively increase the level of detail and concreteness for the actions needed to satisfy the desire. Finally, the final net (the net described in the basic model) generates the neuromuscular activations required to perform the selected action(s).

Essentially, we suggest that in the nervous system voluntary actions are triggered by formulating their end-effects as high-level sensory representations of the desired results. This model is consistent with increasing

evidence from motor research in primates and humans (reviews in Lebedev and Wise, 2002; Graziano, 2006; Cisek and Kalaska, 2010).

4.5 Conclusions

This unsupervised ANN simulation confirmed, as the IMP claims, that voluntary actions can be initiated by imagining (desiring) their sensory effects. IMP seems a valid model for understanding human sensorimotor mapping, intentional movement and motor learning. Detailing and extending the IMP in what we termed the 'chained triplet-net' model makes the IMP also helpful in explaining voluntary behavior besides elementary actions. We are going to add further touches to this picture with the next experiment.

5 'ACON' (ARTIFICIAL CONSCIOUSNESS) SIMULATION

5.1 Introduction

In the ARM model, Motor Will generated random sensory-coded target points where the triplet-net would have to move the hand, and the triplet-net only performed sensorimotor translations needed to execute those motor commands. In this new simulation we extend the ARM model to show how a Motor Will network and a triplet-net together, when provided with a few additional elements and connections, acquire the ability to produce useful intentional behaviour once they are nestled in an environment promoting such a behaviour.

A straightforward definition for a voluntary action is 'an action taken after foreseeing its effects' (Parisi, 2006). The ability to foresee relies on memories from past experience. Hypothesizing that memories in brain are contained in the associative areas and can be recalled by other areas (namely the prefrontal areas), we decided to verify whether evolutionary selection can produce a neural network able to trigger recalls in an associative neural network memory and use them to foresee possible action outcomes, so as to ultimately take useful actions. Even though very simplified, this represents an innovative basic model for brain intentional behaviour.

5.2 Methods

5.2.1 Overall Model Design

This model consisted of a memory containing knowledge about the environment, and another component that used that knowledge to establish which actions to take and therefore acts as a Motor Will. The memory component was a triplet-net enriched with two-way connections between its three unit groups, so as to actually become an associative neural net memory, or auto-associator (McClelland and Rumelhart, 1986) rather than 'just' being a sensorimotor translator like it was in the ARM simulation. Also, Motor Will was not just a generator for random commands like it was in the ARM simulation, but a fully featured ANN sending the triplet-net meaningful activation sequences driving it to recall memories and make predictions. Since these Motor Will tasks recalled some of those called 'executive functions' in psychology, in this model we preferred to call it the *executive-net* rather than Motor Will.

We let the executive-net get shaped and made able to drive the triplet-net memory by evolution (genotypic learning), through genetic algorithms selecting numbers and features for the network units and connections. To this purpose we set up 100 agents, differing in their executive-net units and connections, performing actions in an environment where agents able to recall past experiences and foresee action outcomes would produce more beneficial behaviour, because there was always a better action, depending on which environment location the agent was. Any action was either affecting the agent fitness score, or making the agent change its location. Each agent was given a memory on the effects of each action in each location, as if the agent had explored the environment in advance. This memory was embedded in the triplet-net associative memory.

In order to verify our assumption that performing useful behaviours in the designed environment would require the ability by the executive-net to drive the triplet-net memory, we also carried out a control simulation with agents having no connections from the executive-net to the triplet-net (sort of a lobotomized population).

Lacking those connections, the executive-net could not drive the triplet-net, even if its functions to act into the environment were preserved.

5.2.2 Environment

The environment (Fig. 9) consisted of 48 simple 'Y' mazes with three rooms each. In each room only three action were possible: taking the left door, taking the right door, or staying still. In one room, that we call the 'introductory room', one door leaded to the 'reward room', the other door to the 'penalty room', with no change on the agent fitness score. In the reward room, taking one door resulted in getting a big fitness reward,

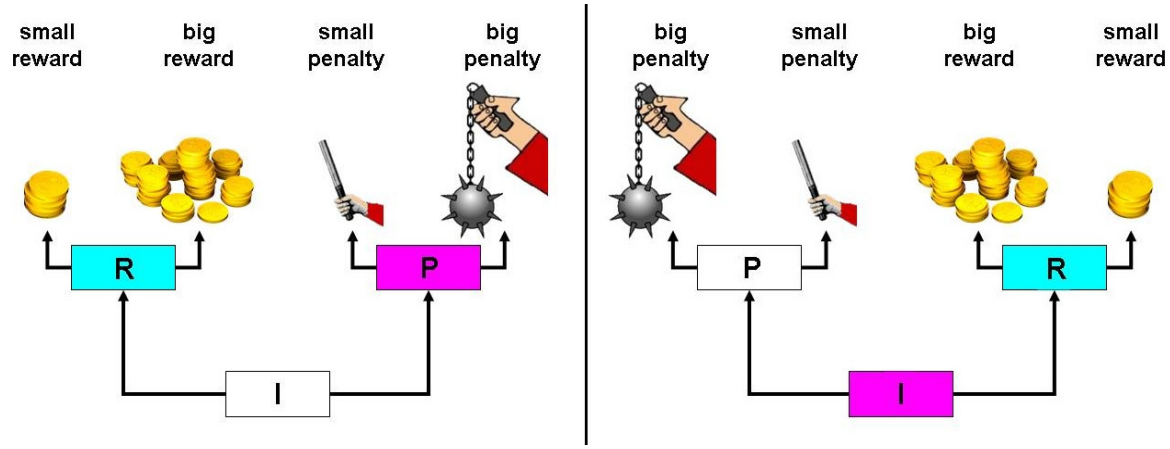


Figure 9: two mazes from the 48 constituting the environment. Each maze comprises three rooms with two doors each. R = reward room; P = penalty room; I = introductory room. Room colors and penalty/reward doors are exchanged in the 48 mazes so that all mazes differ from each other. Each agent is tested in all 48 mazes starting in every room.

the other door a small reward. Finally, in the penalty room one door resulted in getting a big penalty, the other door a small penalty. Reward and penalty doors would not make the agent change room, rather they would make the current maze test end.

Each of the three rooms was colored in a different color, say red white or blue. Each of the 48 mazes was characterized by a unique combination of room colors and door outcomes, and the 48 mazes covered all possible color and door combinations. Any agent could perceive the color of the room in which it was (see also section 5.2.4), but it got no information on whether that room was the introductory or reward or penalty room, until its executive-net would ask its triplet-net about it (by asking about the outcome of taking a door in that room).

5.2.3 Neural Network

The agent neural system (Fig. 10) consisted of an input layer, an

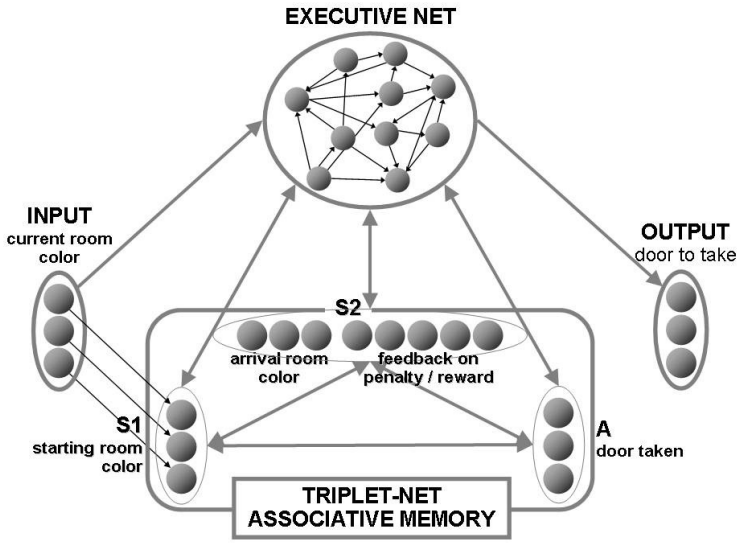


Figure 10: Neural system for the ACON model. Thick grey arrows between unit groups represent full connections between units (any unit in a group sending connections to all units in the other group).

output layer, a triplet-net and an executive-net.

The input layer consisted of three binary neural units encoding the color (blue, red or white) of the room where the agent was, with each unit encoding a possible color as its '1' state. These were simple linear units, just passing their activation values to the triplet-net and executive-net units they were connected to.

The output layer consisted of three analogic neural units encoding the action that the agent would actually perform: activating the first unit would result in the agent taking the left door, the second unit in taking the right door, and the third unit in staying still. The action actually performed was stochastically selected on the difference between the three unit activation values.

The triplet-net was similar to the ARM model, but in this model all its units were binary units, and there were two-way connections between all three unit groups so that the net could work as an associative memory. The three S1 units received information from the input units about the color of the room where the agent was. The S2 units encoded the color of the arrival room, i.e. the room where the agent would arrive because of taking an action (three units like the S1 and input units); they also encoded the fitness reward or penalty related to that action (five units: big reward, small reward, no change, small penalty, and big penalty), for a total of eight S2 units. The motor units (three units like the output units), that in this simulation we prefer to call the 'A' units (A for Action), encoded the action causing to pass from the starting room encoded in the three S1 units (which might differ from the actual room encoded in the input units, when their activations were generated by the executive-net) to the hypothetical arrival room encoded in the first three S2 units. This encoded hypothetical action might differ from the actual action encoded in the output units as commanded by the executive-net. In this model the triplet-net had no connections to the output units, only the executive-net could activate them.

In reality, in this simulation we did not actually implement the triplet-net as a fully featured ANN, but we actually *emulated* it, the same way as in the ARM model we did not actually implement Motor Will, but we rather used a random number generator to emulate it. In this simulation in place of triplet-net actual connections we used a table containing 0 and 1 values emulating the 0-and-1 vectors actually present in the environment, the same vectors that a real triplet-net would learn by experience in the environment. Anytime the executive-net would send the triplet-net an incomplete or wrong vector, the best matching actual vector was searched through the table and returned to the executive-net, thus emulating an associative neural network memory that would fill in or replace all missing or wrong values in the activation vector.

Finally, the executive-net was a non-layered neural network, a set of neural units each of which received connections from the input units, the triplet-net units, and other executive-net units, and sent connections to the output units, the triplet-net units, and other executive-net units. The number of units in the executive-net ranged from 24 to 48 and was determined by evolutionary selection, as also those unit biases, connections and connection weights.

5.2.4 Simulation Flow

When the simulation began, 100 agents were initialized by giving each of them a random number of executive-net units as described in the previous section, and connection weights and unit biases initialized with random values ranging from -0.25 to +0.25. After the initialization stage, the simulation proceeded in turns and generations, with each generation composed by a variable number of turns depending on agent actions.

In each generation all 100 agents were tested in all 48 mazes, starting three times in each of the three rooms for a total number of nine tests in each maze (we chose to perform three tests in each room in order to reduce chance effects given by stochastic processes in the output units, see section 2.3). Agents were tested individually. Before entering a maze, each agent was provided with an emulated triplet-net memory suited to that specific maze, and all unit activations were reset. Each test then proceeded in turns.

In each turn the agent input units perceived the room color where the agent was, and passed it to the triplet-net S1 units and to the executive-net units. The executive-net was activated by it and sent the triplet-net an activation vector overwriting the current S1 activation values from the input units. The triplet-net modified or completed the vector so that it would respect a color (room) + action (door) + outcome combination actually present in the current maze, and sent the result back to the executive-net. On this 'answer' the executive-net 'decided' which action to perform (which door to take, or staying still) and consequently activated the output units. Finally, the action now coded in the output units was executed and its effects (room or fitness score change) applied. All actions were recorded for later evaluation offline. If an agent would stay still for more than

two turns in the reward or penalty rooms, or four turns in the introductory room, he was given a small penalty, the test was ended, and a new test was started.

Once all agents had performed all tests, a generation was ended and the agents were either reproduced into the next generation or discharged, according with their final fitness score. In case of tied scores preference was given to the faster acting agent(s). During reproduction, point mutation and genetic crossover were randomly applied to the offspring, with bit mutation rate = 0.0001 and crossover rate = 1. The main simulation was ended when the correct actions in any mazes and any rooms would exceed 95%, or after 1000 generations with no improvement > 1% in any maze or any room. In order to properly compare the experimental population with the control population, control population simulation was carried on up to the same generation number as the experimental population even after no substantial improvement occurred along over 1000 generations.

5.2.5 Mind Reading

After the simulation ended, the overall best agent in the last generation was tested in all mazes once again, and interactions between its executive-net and triplet net were analyzed to uncover the agent 'reasoning' strategy. This was possible because those interactions consisted in activation vectors on the triplet-net units reflecting the vectors actually present in the environment, vectors of which we knew the meaning of each 0 or 1.

5.3 Results

After about 13000 generations the agent population reached 99% correct ('good') actions in all rooms in all mazes (Fig. 11 upper row). Reducing the mutation rate to one tenth of the original rate led in a further 1000

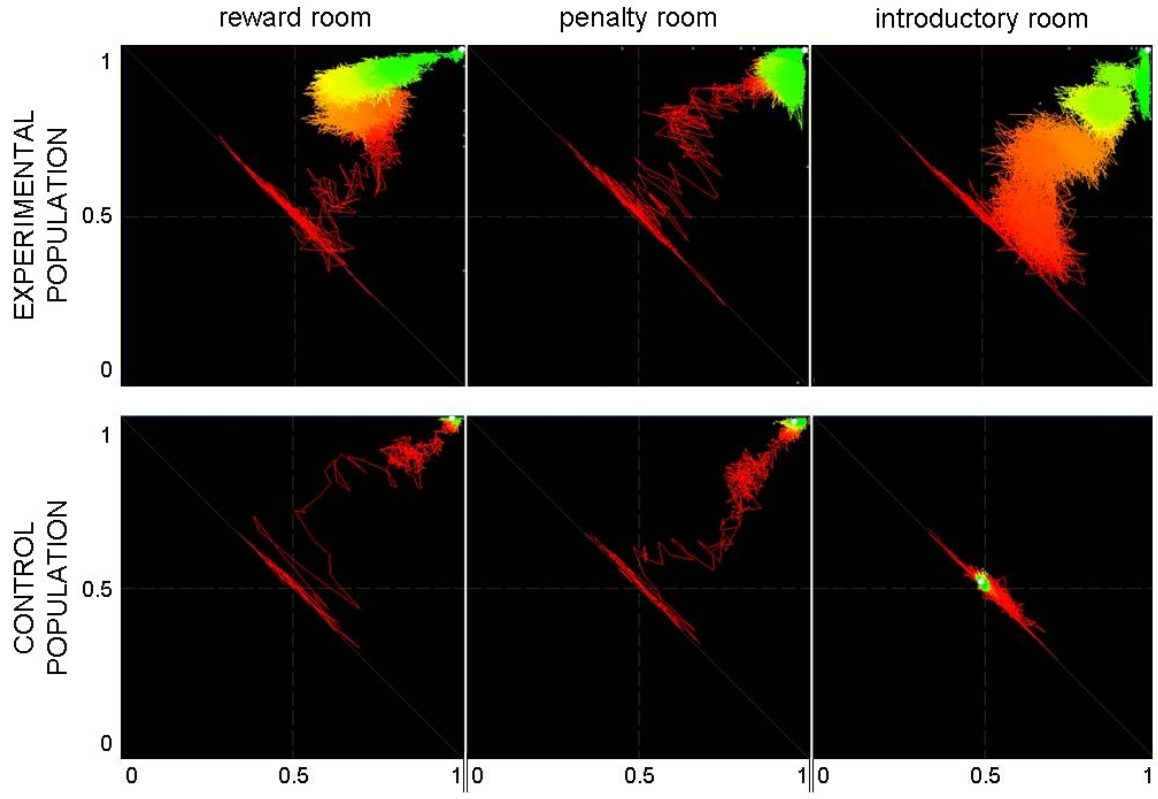


Figure 11: Evolutionary path for 13000 generations, measured as the mean prevalence of correct actions performed by agent population in the reward room (left), penalty room (middle) and introductory room (right). Vertical axis: values for the left door; horizontal axis: values for the right door. Upper row: experimental population; bottom row: control population with no connections from the executive-net to the triplet-net. Each coloured point represents one generation, with red colour for earlier generations and green colour for later generations. Each window center point represents 50% correct actions, that is altogether random actions; the top-right corner represents 100% correct actions for both the left and right door, i.e. evolutionary success. Experimental population was successful in all three rooms. Control population was successful in the reward and penalty room, solvable by automatic action; but it never deviated from random actions in the introductory room, where taking the correct door needed imagining what would happen next.

generations to almost 100% correct actions. The executive-net consisted of 30 units in every agent.

The control population (Fig. 11 bottom row) reached over 95% correct actions in the reward or penalty room after just 500 generations, but it never deviated from 50% (corresponding to all random results) in the introductory room, even after 13000 generations. The executive-net consisted of just 3 units in every agent.

The 'reasoning' strategy developed by evolution in the successful population consisted in the executive-net activating the triplet-net with a same sequence of questions in every maze. Of course those questions were just vectors with 0 and 1 values; here we translate in common human language what those 0 and 1 meant.

a) On the first turn, the first question invariably was: 'what happens if in my current room I take the left door?' If the triplet-net returned a reward or penalty, thereby notifying the agent of being in the reward or penalty room, then the executive-net immediately knew which door to take (the same left door if the triplet-net returned the big reward or the small penalty, the opposite door if the triplet-net returned the small reward or the big penalty), and it made the agent acting accordingly.
b) If the triplet-net instead returned a different room and no reward or penalty, thereby notifying the agent of being in the introductory room, then the executive-net made the agent staying still, and on the second turn it asked the triplet-net a second question: 'what is going to happen if in the room beyond the right door I will take the left door?' If the triplet-net reported a reward, either a big or small one, now the executive-net had all information needed to go to the big reward (it would take the right door from the introductory room to the reward room, then it would take the door identified by the second answer as the big reward door).
c) If otherwise the second answer reported a penalty, then the executive-net commanded the agent to take the other door (i.e. the left door), thereby going to the reward room, and on the third turn it asked the triplet-net a third question: 'what happens if in this room I take the left door?' This was like the question in a) and it worked the same way (see above).

With this strategy, agents could solve every maze in just one turn when starting in the reward or penalty room; three turns when starting in the introductory room. This was one turn less than the maximum allowed.

5.4 Discussion

In designing this simulation we chose to emulate triplet-net memories rather than use full-featured ANN memories, because we needed working memories deprived of their learning ability. Even if this was self-contradictory and very different from real nervous systems, we needed it because in real nervous systems repeated experience leads to automatic behaviour, and this occurs because nervous systems can learn. Automatic behaviour is the opposite of intentional behaviour, and since this experiment was focused on intentional behaviour we had to prevent our agents from *learning to act*; we wanted our agents *learn to think* what actions to perform.

We solved this problem by allowing only genetic learning in the executive-nets, and no learning in the triplet-nets. The triplet-nets could not learn because they were emulated nets, with no modifiable internal connections. As of the executive-nets, genetic learning would not let them develop automatic responses, because genetic learning can only capture constant environmental features, but the designed environment contained no constant features, since the 48 mazes that any agent experienced were altogether different in their features. Therefore the only abilities that genetic learning could provide to the executive-nets were the ability to drive the triplet-net with a same sequence of activation vectors useful in any maze, and the ability to select appropriate actions on the answers delivered by the triplet-net.

Our results unequivocally showed that the proposed neural architecture, consisting of an associative memory net containing information on the environment, connected to an executive-net triggering memory recall in the memory net and using those information to achieve goals, is a working model for the way by which nervous systems can generate useful, intentional actions. As we said in the introduction, a voluntary action is an action taken after foreseeing its effects. In the present model, foreseeing action effects is performed by the triplet-net on a hypothesis generated by the executive-net. When a triplet-net completes or modifies the S2 activation values in an activation vector, it is actually making a prediction about the outcome of taking the door encoded in the A units when being in the room encoded in the S1 units – a door and a room generated ('imagined', see below) by the executive-net by properly activating the A and S1 units.

In this model we used a triplet-net as the associative memory because a triplet-net is ideal to encode relations between actions and their effects. External activations to the triplet-net may come from sensory input units as well as from executive-net units. Activations coming from input units can only recall memories in

which the starting room is the same as the actual current room, which makes sort of an 'automatic' recalling. On the contrary, activations from the executive-net can also evoke memories in which the starting room differs from the current room provided by the sensory input, which is *imagining*. In the designed environment, reward and penalty rooms can be solved with automatic recalling. On the contrary, the introductory room requires two or three successive recalls, the second of which (cf. section 5.3 step b) always concerns a different room than the room currently reported by sensory input, i.e. imagination of being in another room. Both the ask sequence revealed by 'reading the mind' in the successful population, and the fact that the control population reached success in the reward and penalty rooms, not in the introductory room, attest that this more abstract imagination depends on the executive-net, even if it occurs in the memory net.

Insofar as the agents performed actions after foreseeing those action effects, according to Parisi (2006) we can assume that they performed voluntary actions. Still, our agents lack a relevant feature typical of the human will: the ability to change one's own goals. Our agents' one and only goal invariably was to increase their fitness score, a goal not even made explicit in their 'mind' (i.e. not represented as unit activations), but rather embedded by genotypic learning into their executive-net units and connections. This condition with one or a few built-in, unchangeable goals, that the agent is even unaware of, resembles that of animals or artificial intelligence chess playing programs, rather than that of man. Also, imagination in these agents only consists in recalling memories on past experience and using them, these agents cannot imagine anything they did not experience in the environment, neither they can create innovative solutions. Developing agent able to choose their own goals and to produce creative thinking are aims left to future research.

6 CONCLUSIONS

Evolutionary connectionism is a feasible and profitable methodology to study nervous system functions at any level, from low-level sensorimotor functions up to high-level cognitive functions. Connectionist models are supported by simulation (actually, they *consist in* simulations), which makes them something more than mere theories, and even if they are necessarily simplified with respect to their biological counterparts, they perform well in highlighting essential processes for the investigated functions and the way they emerge in biological nervous systems. We described a connectionist model demonstrating what IMP had been suggesting for over a century, i.e. that voluntary actions are initiated by imagining their sensory effects. Detailing and extending this model to what we termed the 'triplet model' led us to formulate an original model for cognitive functions such as imagining, foreseeing, and ultimately producing voluntary behavior. Our results with these models make us confident that future research by evolutionary connectionism might help to explain even higher functions, such as the faculty to choose one's own goals, up to ultimately the very feeling of being conscious.

REFERENCES

Accornero, N., and Capozza, M., 2009. *Coscienza Artificiale*, Aracne, 303-412.

Bizzi, E., 1974. The coordination of eye-head movements. *Scientific American, 231*, 100-106.

Calabretta, R., and Parisi, D., 2005. Evolutionary Connectionism and Mind/Brain Modularity. In Callebaut, W. and Rasskin-Gutman, D. (eds.) *Modularity. Understanding the development and evolution of complex natural systems*. The MIT Press, Cambridge, MA, 309-330.

Caudill, M., and Butler, C., 1992. *Naturally Intelligent Systems,* MIT press.

Chalmers, D. J., 1990. The evolution of learning: An experiment in genetic connectionism. In Touretzky, D.S., Elman, J. L., Sejnowski, T. J. and Hinton, G. E. (eds.) *Proceedings of the 1990 connectionist models summer school*, Morgan Kaufmann, San Mateo, CA, 81-90.

Cisek, P., and Kalaska, J. F., 2010. Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, *33*, 269-298.

Flanagan, J. R., and Rao, A. K., 1995. Trajectory adaptation to a nonlinear visuomotor transformation: evidence of motion planning in visually perceived space. *Journal of Neurophysiology*, *74*(5).

Goldberg, D.E., 1989. *Genetic algorithms in search, optimization, and machine learning.* Addison-Wesley Publishing Company, Inc., Reading, MA.

Graziano, M., 2006. The organization of behavioral repertoire in motor cortex. *Annual Review of Neuroscience*, *29*, 105-134.

James, W., 1890. *The principles of psychology*, Harvard University Press, Cambridge, MA.

Kiesel, A., and Hoffmann, J., 2004. Variable action effects: Response control by context-specific effect anticipations. *Psychological Research*, *68*(2-3), 155-162.

Lebedev, M. A., and Wise, S. P., 2002. Insights into seeing and grasping: distinguishing the neural correlates of perception and action. *Behavioral and Cognitive Neuroscience Reviews*, *1*(2), 108-129.

McClelland, J. L., and Rumelhart, D. E., 1986. A distributed model of human learning and memory. In Rumelhart, D. E., McClelland, J. L., and The PDP Research Group, *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, Vol. 2, MIT Press, 209-214.

Parisi, D., 2006. *Una nuova mente*, Il Mulino.

Rumelhart, D. E., Hinton, G. E., and McClelland, J. L., 1986. A general framework for parallel distributed processing. In Rumelhart, D. E., McClelland, J. L., and The PDP Research Group, *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, 45-76.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986. Learning Internal Representations by Error Propagation. In Rumelhart, D. E., McClelland, J. L., and The PDP Research Group, *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, 318-362.

Schaffer, J. D., Whitley, D., and Eshelman, L. J., 1992. Combinations of genetic algorithms and neural networks: a survey of the state of the art. In *International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN-92)*, Baltimore, USA, 1–37.

Shadmehr, R., 2005. *The computational neurobiology of reaching and pointing: a foundation for motor learning*. MIT press.

Shadmehr, R., and Arbib, M. A., 1992. A mathematical analysis of the force-stiffness characteristics of muscles in control of a single joint system. *Biological cybernetics*, *66*(6), 463-477.

Stock, A., and Stock, C., 2004. A short history of ideo-motor action. *Psychological Research*, *68*(2-3), 176-188.

Yao, X., 1993. A review of evolutionary artificial neural networks. *International journal of intelligent systems*, *8*(4), 539-567.

Yao, X., 1995. Evolutionary artificial neural networks. *Encyclopedia of computer science and technology*, *33*, 137-170.

Yao, X., 1999. Evolving artificial neural networks. *Proceedings of the IEEE*, *87*(9), 1423-1447.

CHANGING SHAPES AND FUNCTIONS IN CELL POPULATIONS SIMULATED BY MULTI AGENT SYSTEMS.

Alfredo Colosimo
Dept. S.A.I.M.L.A.L., Sapienza University, Via A. Borelli 50, 00161 Rome, Italy
alfredo.colosimo@uniroma1.it

Keywords: Multi Agent Systems (MAS), Neural Networks, Phenotypic Changes, Cells and Force Fields.

Abstract: The simulation strategies based upon Multi Agent Systems (M.A.S.), taken together with the flexibility of some of the associated programming environments, show an outstanding heuristic power. In particular, the realistic reproduction of several complex phenomena characteristic of biological systems is made possible by means of the same, unifying approach. In this frame, we simulate: a) the phenotypic changes in a generic cell population under the influence of endogenous and exogenous factors, and b) the performance of small artificial neural networks, mimicking some highly simplified information processing mechanisms. In both cases we focussed on the qualitative reproduction of natural time-courses.

Contents

1 INTRODUCTION

Among the multifaceted strategies to tackle the complex problems arising if one tries predicting the collective behavior of a cellular population on the basis of the information concerning its elements, Multi Agents Systems (M.A.S.) are endowed with the following desirable features:(DeToni and Bernardi, 2009), (Dzitac and Barbat, 2009), (Ferber, 1999) :

- a flexible simulation procedure which can be used at quite different dimensions, from the microscopic (single cells) to the macroscopic (tissues, organs) scale;
- some powerful software tools like NETLOGO (Wilenski, 1999), stemming from well established programming environments typical of Artificial Intelligence (Russell and Norvig, 1995), such as PROLOG and StarLogo (http://education.mit.edu/drupal/starlogo-tng).

The systematic exploitation of this approach in the study of phenomena concerning different disciplines, from Sociology and Economy to Neuroscience and Cell Biology, produced relevant contributions whose quality can be grasped by looking at the NETLOGO (Wilenski, 1999) models library. In what follows we illustrate how we took advantage of that in handling two of my research topics at CISB in recent years, namely:

- the influence of environmental force fields (e.g. electromagnetic or gravitational) on the morphofunctional phenotypic changes occurring in cellular populations (Bizzarri et al., 2011);
- the information processing mechanisms in artificial neural networks and their dependence upon the (a)symmetric arrangement of links (Holland, 1992), (Holland, 1996).

In the former case, changes in the relative abundance of interconverting cellular populations were made sensitive to aggregation phenomena induced by external force fields. In the latter case, the influence of (randomic) rearrangements of connections between the network nodes was quantitatively investigated in model systems of minimal size by means of a new and highly flexible procedure of data input in the M.A.S. simulator.

All simulation were carried out by means of the NETLOGO tool (Wilenski, 1999) and extensive use of its tutorials and ready-to-use macros. Reproduction and extension of our results might be facilitated by the Supplementary Material in the Appendix and, whenever necessary, by specific inquiries to the author (indicating "CISB Meeting 2014, M.A.S.2" in the e-mail subject).

2 METHODS

2.1 influence of intrinsic and extrinsic factors on cellular phenotypes.

According to a crude but very useful approximation, in a cellular population including two subgroups, (C_1) and (C_2), the rate of switching, (v_1), from (C_1) to (C_2) can be associated to a single first-order process, ($C_1 -> C_2$), ruled by a kinetic constant k_1. This leads to: $v_1 = k_1 * [C_1]$, where the square brackets indicate molar concentration. If the switch is reversible, the same reasoning applies to the reverse process, $C_2 -> C_1$, leading to: $v_2 = k_2 * [C_2]$. At equilibrium v_1 and v_2 have the same value; thus, $v_1 = k_1 * [C_1] = v_2 = k_2 * [C_2]$, and the relative proportion of C_1 to C_2 is given by: k_2 / k_1. A basic M.A.S. model of the process, including a preliminary description of the force field effects, as summarized in Figure 1, can be found in a previous work (Colosimo, 2011) .

The new model discussed here includes the sensitivity of kinetic constants to external force fields. In the standard formalism for a first-order kinetic process described by the differential equation $(dA/dt) = -k^{'} *$ A, the time course of the exponential decay of A_0 is given by:

$$A_t = A_0 * exp(-k^{'} * t) \tag{1}$$

A straightforward account for any external influence on the rate of that reaction, would obviously invoke the role of a generic Environmental Factor (EF) as an additive factor associated to the intrinsic constant k :

$$k^{'} = k + EF \tag{2}$$

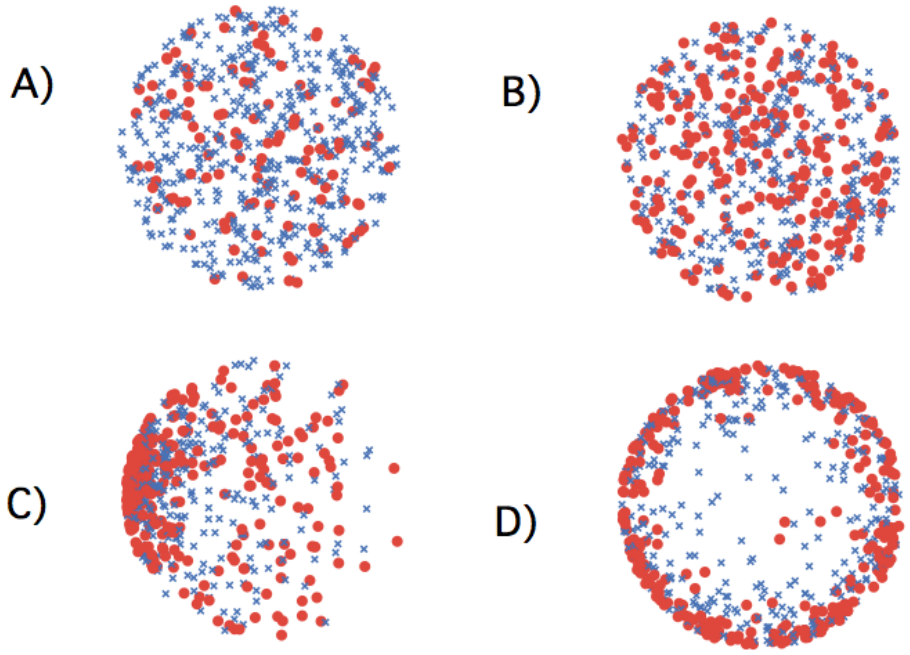


Figure 1: **Environmental force fields and cellular phenotypes.** In **A)** and **B)** the yellow/blue cells ratios are 1/4 and 3/4, respectively, as defined by the interconversion rate constants (see the text). In **C)** and **D)** the ratio is 1. However, in **C)** the yellow cells are concentrated in a small region on the left of the Petri dish under the action of a specific attractive field; in **D)** a repulsive field, specific for the blue cells, is located in the centre of the dish.

In the special case of Gravitation (G) being the dominant external force field, some further detail could be easily considered by such a model as, in particular, the influence on the interconversion rate of the attraction exerted by G on the cell population. In other words, the kinetic constant could be modified by the relative increase of cellular density in a given space region, making the size of EF factor in Eq. 2 directly proportional to the intensity of the force field at hand.

All in all, the overall effect of the force field can be made:

- a) attractive or repulsive: the latter condition could mimic a physical (e.g. temperature) or chemical (e.g. drug) unfavourable gradient;
- b) of decreasing intensity from the source location according to a given law: we choose here a simple inverse dependence from the source distance.

Figure 1 shows how the results of a M.A.S. dynamic simulation can be visualized in the form of events occurring in a Petri dish, looking at time dependent changes and spatial distribution of phenotype shapes .

The relative amounts of Yellow and Blue phenotypes at equilibrium, in the Petri dish (**A** and **B** panels of Figure 1), provide the same information as that in Figure 2 under equilibrium conditions (as defined at the beginning of the kinetic regimes (I) and (III). However, it should be noticed that in all the four panels of Figure 1 the ratio of kinetic constants is 1: since no influence on the kinetic constants by force fields has been taken into account there, the apparently different number of cells in the **C)** and **D)** panels is an optical effect solely due to the higher density of yellow cells in a small area of the **C)** panel.

The effective influence of force fields on interconversion of phenotypes, is described in Figure 4.

2.2 Basic topological network descriptors.

Figure 3 aims to illustrate by very simple examples the nomenclature used in this paper, which is also summarized in the Table in Appendix 2. As for the *Average Path Length (APL)*, notice that the Average Path Length (identical to the mean Shortest Path Length, $< SPL >$), is a 'global' descriptor concerning the whole network. On the other hand, $SPL_{i,j}$ is a 'local' descriptor providing information over a limited area in the network, since it concerns the shortest path connecting the specified couple of nodes only., The relation between the pattern of links and APL, given by the formula:

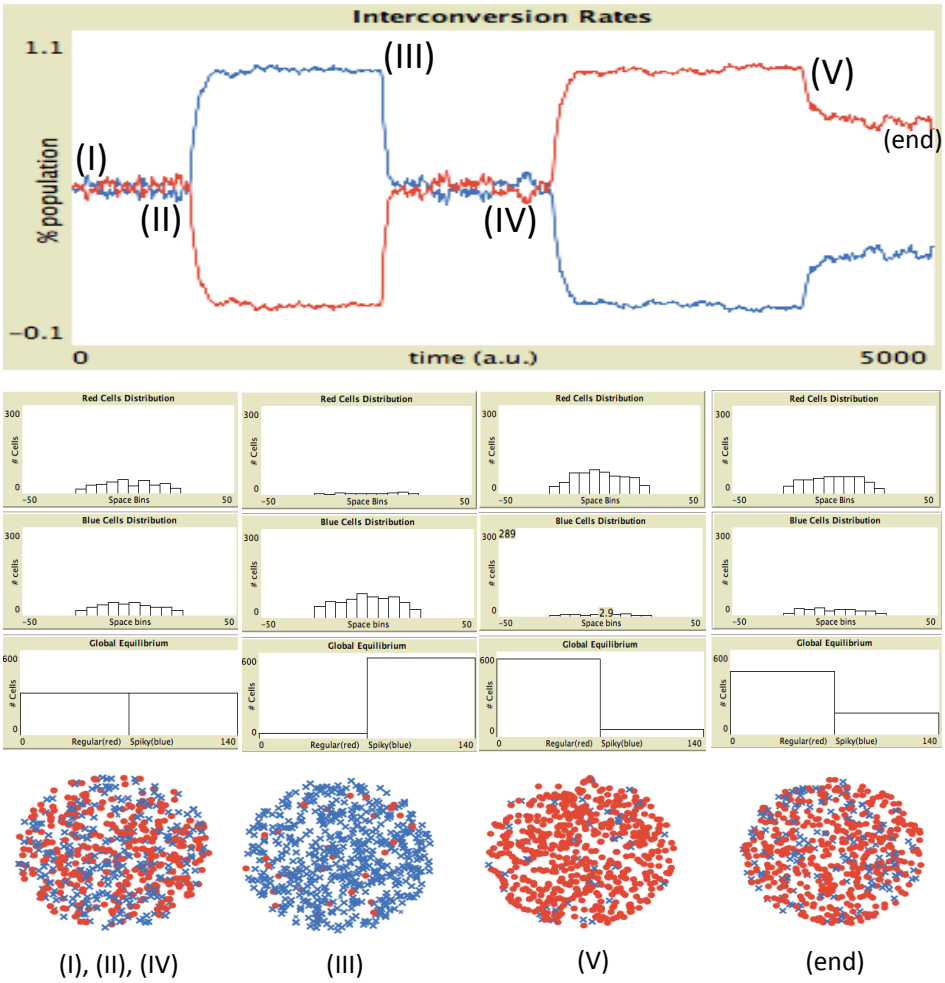


Figure 2: **Interconversion between a Red and a Blue cellular subpopulation in different kinetic regimes.**
Upper panel: The (I) - (V) symbols indicate the start of the corresponding regime, defined by the k_R and k_B kinetic constants: $R-(k_R)->B$, and $R<-(k_B)-B$. The kinetic regimes are defined by the ratio k_R/k_B as follows: 1 (I), (III); 20 (II); $1/20$ (IV) and $1/3$ (V). Time is in machine units; the number of cells at t = 0 is: Red = Blue = 300. The equilibrium is reached when the average oscillations remain within 1% in a time window of ten times the reciprocal of the slower kinetic constant.
Lower panel: The spatial distribution of Red and Blue cells in the Petri dish at the time points I -IV (and final) is indicated in the form of histograms (Middle) and directly visualized (Bottom).

$$APL = \frac{1}{N(N-1)} \sum_{i \neq j} d_{i,j} \tag{3}$$

where N = number of nodes and $d_{i,j}$ is the shortest number of links to be passed through in going from node i

to node j provides, as an easily-predictable outcome the unitary value associated to fully connected networks of any size (see panels **C)** and **E)** of figure 3.

Comparing the regular link arrangement in panels **A)** and **D)** of figure 3, where only the first neighbors of each node are involved, two interesting observations concern: a) the higher APL value for the bigger net (2.5 as compared to 1.5), and b) the same 'null' value for the average clustering coefficient (see below).

While an appropriate reckoning of APL is only carried out by a dynamic programming algorithm [1], in spite of its simple physical meaning, on the contrary, calculation of the *Clustering Coefficient (cc)*, is relatively easy, but it deserves a little more attention. The *Clustering Coefficient (cc)*, is defined as the ratio of existing links connecting a node's first neighbors to each other, to the maximum possible number of such links. Thus, for the i_{th} node this is given by cc_i (see the Table in Appendix 2) which provides the number of "triangles" that go through that node, whereas $\frac{K_i(K_i-1)}{2}$ is the total number of triangles that could pass through node i, should all of node i's neighbours be connected to each other. The clustering coefficient for an entire network, $< cc >$, is the average of the clustering coefficients of all the nodes namely:

$$< cc > = \frac{1}{n}\sum_{i=1}^{n}\frac{2(AT)_i}{K_i(K_i-1)} \tag{4}$$

where N is the number of nodes and $(AT)_i$ (the number of the actual triangles including node i and its first neighbors), corresponding to e_i in Table 3.2. It is maybe worth reminding that a relatively high $< cc >$, taken together with a relatively low value of APL, is an indication of the *small world* architecture of the network. [2]

...

3 RESULTS AND DISCUSSION.

3.1 Force fields influence cellular phenotypes.

The main purpose in simulating the time course of interconversion between cellular phenotypes was to reproduce the effect of modified cellular densities induced by external forces in specific regions of the environment. To this aim we modulate the kinetic constants describing in terms of simple first order kinetics the interconversion processes (Eq 1), by introducing a modifier (Eq 2) which, in turn, may depend upon local cell densities in specific regions of the Petri dish. More specifically, the generic Environmental Field (EF) of Eq. 1 is now defined as the product of two quantifiers of cell overcrowding, named "Overcrowding" and "InRadius". The former term is a purely empirical modulator, while the "InRadius" term is more interesting since it measures the number of cells of the same type close to each other. This number may continuously change due to the attracting/repelling effect of the force field. In summary, the kinetic constant for each of the two interconversion process is given by:

$$k^{'} = k + Overcrowding * InRadius \tag{5}$$

where: InRadius is the number of cells whose distance from the reference cell is below a given threshold [3].

It is worth reminding that in our M.A.S. simulations cells are continuously moving in their environment and may distribute following one or more force fields, whose intensity is inversely proportional to the distance from the source. As shown by Figure 4, where the two panels refer to a variety of kinetic parameters values, this approach showed of noticeable flexibility and effectiveness. The set of original NETLOGO routines use to work out all the simulations of cell phenotype changes is reported in Appendix 1.

[1] see the Floyd-Warshall algorithm: http://en.algoritmy.net/article/45708/Floyd-Warshall-algorithm

[2] The term *small world network* originated in the social sciences many years ago now (Milgram, 1967) to reflect the assertion that within networks of social acquaintances all people on the planet are separated from each other by just a small number of intermediate acquaintances. More recently, it was shown (Barabási and Oltvai, 2004) that the distance between any couple of individuals, d_{ind}, measured by the intermediate individuals or intermediate links, is much smaller than expected. Moreover, The famous "six degrees of separation" estimated at that time, nowadays are probably decreased, thanks to the WEB and the Social Networks.

[3] The modulation of the kinetic constant could be of a different form like, for instance, $k^{'} = k * (Overcrowding * InRadius)$.

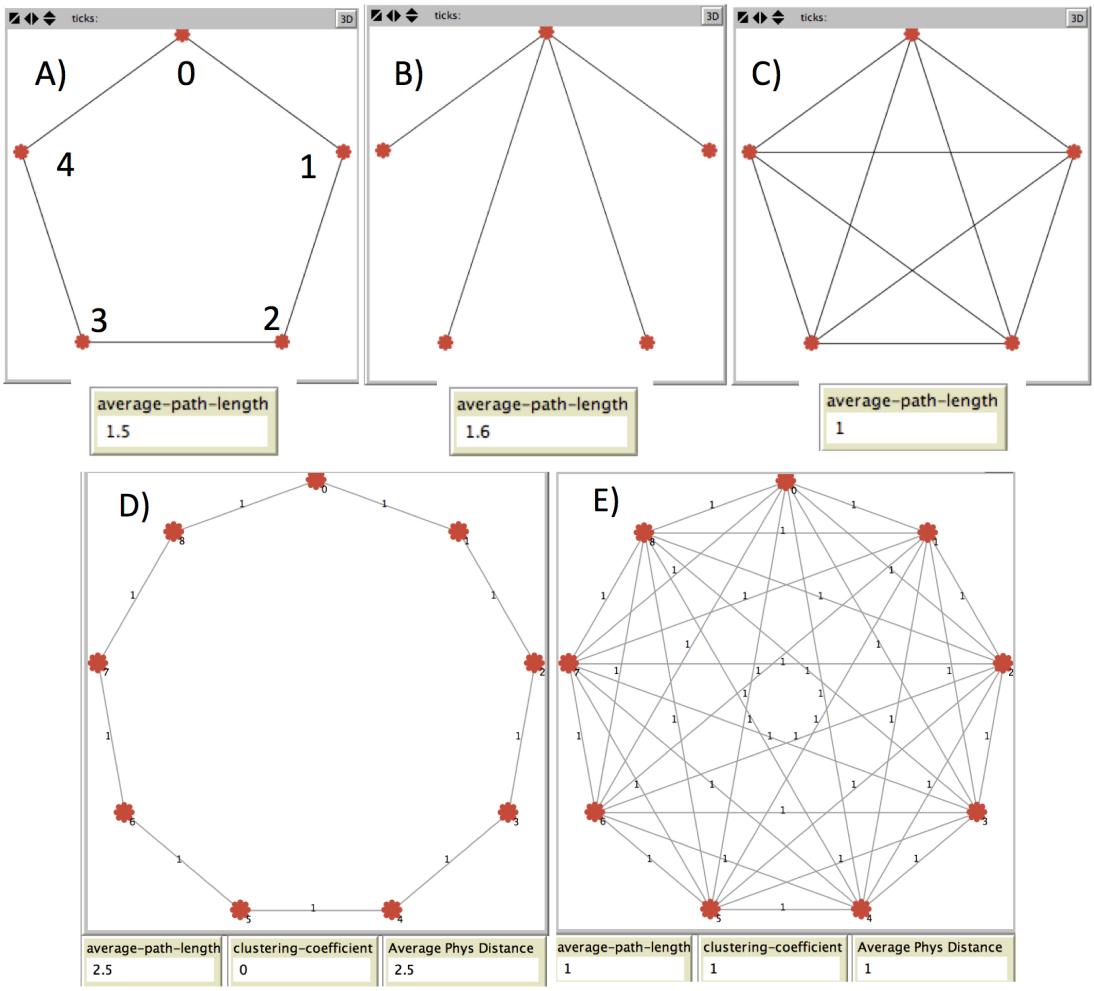


Figure 3: **Different patterns of links in simple networks.**
The rule followed in the A) panel to identify nodes by numbers applies to all other panels.
Top Row (5-node net): The Node Degree estimates for each node (in the order dictated by increasing numbers) are the following: A) = 2, 2, 2, 2, 2; B) = 4,1,1,1,1; C) = 4,4,4,4,4. Notice the different APL values in the three panels.
Bottom Row (9-node net): The Average Path Length values (see the Table in Appendix 2) are flanked by $< cc >$ and Average Physical Distance values, at the bottom of each panel. Notice in the D and E panels the null and unitary $< cc >$ values which account, respectively, for a regular pattern of connections between first neighbors only, and fully connected networks. For details on reckoning APL values, see the text.

3.2 Link patterns influence the efficiency of neural network .

The connections between nodes in a network model endowed with even a minimal biological realism should be based on both spatial (structural/anatomical) and logical (functional/physiological) relationships between nodes. In this frame it is difficult to overestimate the importance of: a) representing such relationships in a matrix form, and b) using a robust procedure for both data insertion and updating.

An example of the matrices we set up for our work is the following, which refers the network in Figure 5 :

```
                Adiacency  Matrix                    'Physical'  and  'Topological'
                                                        distances between nodes

(node 0):  [[ 0  1  0  1  0  0  0  0  0 ]          [0  1  3  1  3  3  2  2  3]
(node 1):   [ 1  0  0  1  0  0  0  0  0 ]          [1  0  3  1  3  3  2  2  3]
```

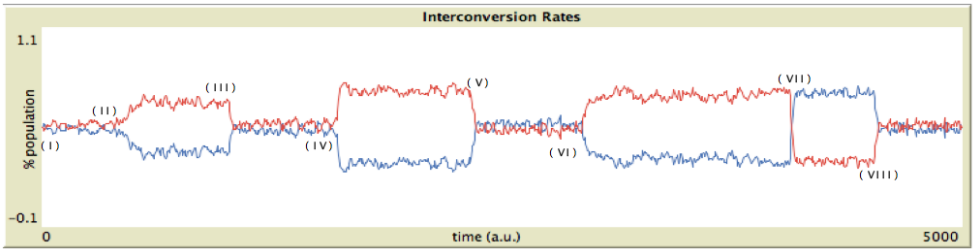
```
(node 2):   [ 0  0  0  0  0  1  0  1  0 ]          [3  3  0  2  2  1  2  1  3]
(node 3):   [ 1  1  0  0  0  0  1  1  0 ]          [1  1  2  0  2  2  1  1  2]
(node 4):   [ 0  0  0  0  0  0  0  1  0 ]          [3  3  2  2  0  2  2  1  3]
(node 5):   [ 0  0  1  0  0  0  0  1  0 ]          [3  3  1  2  2  0  2  1  3]
(node 6):   [ 0  0  0  1  0  0  0  1  1 ]          [2  2  2  1  2  2  0  1  1]
(node 7):   [ 0  0  1  1  1  1  1  0  0 ]          [2  2  1  1  1  1  1  0  2]
(node 8):   [ 0  0  0  0  0  0  1  0  0 ]]         [3  3  3  2  3  3  1  2  0]

In both matrices,   values in the  i,j location concern  nodes  in row i and column j.
```

The Adjacency Matrix (on the left) indicates for each node the links to any other node, namely a full description of the topological structure of the network. The matrix on the right, however, indicates that "Physical" and "Topological" distances between nodes have the same value, which actually occurs only in the absence of any "weighted" link (see below).

Panel **C)** of Figure 5 shows the *Shortest Path Length (SPL) Matrix*, which includes the $SPL_{i,j}$ values for each couple of nodes. The $SPL_{i,j}$ values can used as indexes of local clustering index and, once averaged over the whole network, provide the APL (a global connectivity index). A glance at Panel B) allows a straightforward calculation of the $< cc >$ of the network, as defined in Table 3.2, according to equations 3 and 4.

Substituting in the Adiacency Matrix real numbers to binomial (1 / 0) values is a powerful enhancement of the heuristic power of network models. As a matter of fact the information on the actual "strength" of links between couples of nodes enriches a purely topological viewpoint by including physical considerations, based on concepts like energy and distance among nodes (Latora and Marchiori, 2001).



	(I)	(II)	(III)	(IV)	(V)	(V)	(VII)	(VIII)
Red to Blue k_intrinsic	20	20	40	20	20	40	20	40
Red Crowd Coefficient	/	/	/	/	1	1	/	1
Red Attractor Intensity	/	/	/	/	1/1	1/1	/	1/1
Blue_to_Red k_intrinsic	20	20	20	40	40	20	40	40
Blue Crowd Coefficient	/	1	1	1	/	/	1	/
Blue Attractor Intensity		1/1	1/1	1/1	/	/	1/1	1/1

Figure 4: **Effects of force fields on cellular phenotypes.**
The time courses in the upper panels are dictated by Eq. 5 and the combination of parameters in the table (lower panel). All other details as in Fig. 2 .

Figure 6 contains an even simpler network than that in Figure 5 (9 links instead of 11), where the classical topological descriptors Average Path Length and cc, are now flanked by a new one, Average Physical Distance, simply calculated by the same criteria as APL, just assuming $d_{i,j}$ in eq. 3 as the smallest sum of link weigths while going from node i to j. As expected, changing the weights of just one out of nine links dramatically changes the Average Physical Distance, leaving unaltered both APL and CC (top row of networks in Figure 6.

In the Left, Middle and Right column panels in the figure the same number of links is characterized by a different distribution among the nodes, and by different sets of weights in the Top and Bottom row panels, as defined in Appendix 2.

The lesson to be leant by Figure 6 is how different (and complementary) is the sensitivity of the two descriptors used here (APL and $< cc >$) to the link distribution over the net and to the link strength. Thus, an appropriate combination of the the two descriptors appears able to account for a large number of different experimental data sets.

4 CONCLUSIONS.

This contribution concerns a couple of problems we faced in the research programs carried out at CISB in recent years, in both of which a simulation strategy based on Multi Agent Systems (M.A.S.) came out precious.

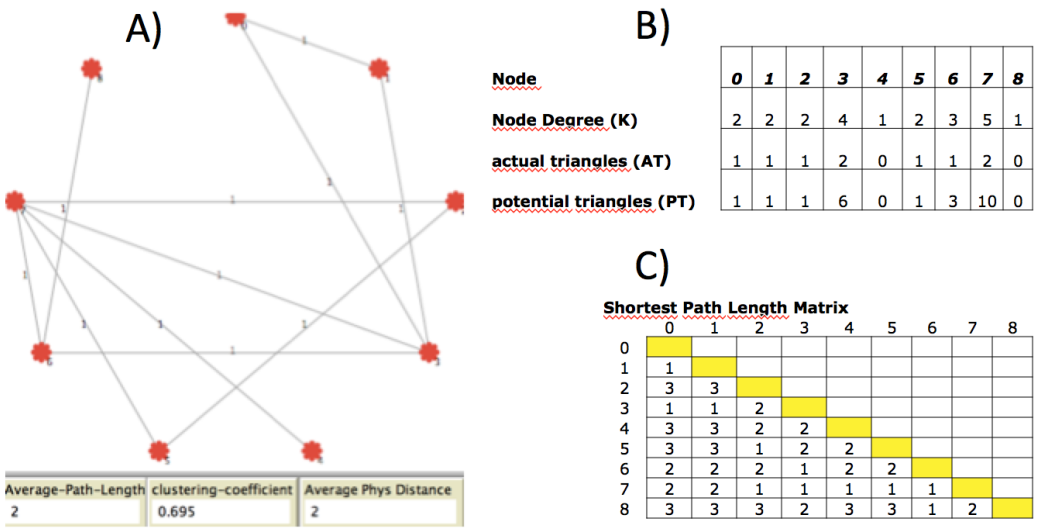


Figure 5: **Calculation of basic network descriptors.**
A): A circular network having 9 nodes and 11 links. Nodes are plotted in clockwise, ascending order from 0 to 8. The links are labelled by their 'strength' (see the text for details).
B): Intermediate values for the calculation of the network descriptors (see the Table in Appendix 2). Average Node Degree = $< K >$ = 2.44; Average Shortest Path Length = $< SPL >$ = 2; Clustering Coefficient = $< cc >$ = 0.695 (see Eq. 1). In the last calculation nodes 5 and 9, having K = 1, are not considered).
C): The Shortest Path Length (SPL) matrix, of the network in A).

In one case, we showed that a M.A.S. paradigm is able to mimic the dynamics of the phenotypic changes induced in cell populations by intrinsic (genetic) as well as extrinsic (environmental) factors. In quite general terms, assuming that the interconversion between morphofunctional phenotypes can be described as a first order reversible process, the associate kinetic constants are easily adjusted by additive/multiplying factors accounting for external force fields of any type. If the global changes at the population level depend upon the vicinity of individual cells to each other, we realized that a M.A.S.-based simulator is very well suited to reproduces changes in cell density in different regions of the available space, consequential to the presence of aggregating (gravitational, electrical, etc.) force fields.

A good performance is revealed by M.A.S. programming also in studying the ordered and/or random global activity of cell populations when interactions between individual cells (nodes) are modeled by specific connections (links) in a well defined network. Thus, another promising evolution of the M.A.S. simulation strategy described here should be in tracing the pattern of functional links responsible for the mutual activation/inhibition of distant regions in brains.[4]. Such a network-like scheme provides a useful, minimal model for the axonal links among brain regions which become synchronously active as a response to several physiological/pathological events. In order to stress the practical advantages of the above described matricial representation of connection links, it is worth noticing that it allows a straightforward consideration of both activating and inhibiting interactions, in the form of, respectively, positive and negative values.

Obvious as it is, the most promising output of the present research line remains the confrontation with reliable experimental data in order to get:

- a mechanistic interpretation of unexpected and apparently counterintuitive observations;
- an intuitive, 'analogical' representation of the complex events arising from the global behavior of interacting units, which is more difficult to obtain through classical (O.D.E.) methods.

...

[4] Trying and reproduce at least some of the peculiar performance of CNS (like learning and memory) by simulating the collective behavior of neurons populations has always been in the dreams of biological modellers (Erdös and Rényi, 1960), (Wolfram, 2002), (Colosimo, 2008). Since the introduction of the 'perceptron' (with the associated back-propagation algorithm) models, or of genetic algorithms (with the associated fitness function) (Accornero and Capozza, 2009), (De Jong, 2002), the most important breakthrough in simulation strategies probably is the *complex network* metaphor (Hegsellman and Flache, 1998). In this metaphor any complex global functioning directly stems from the intricacies of the non linear functional connections (links, edges) ruling the mutual activation/inhibition/synchronization of a high number of functional modules (nodes).

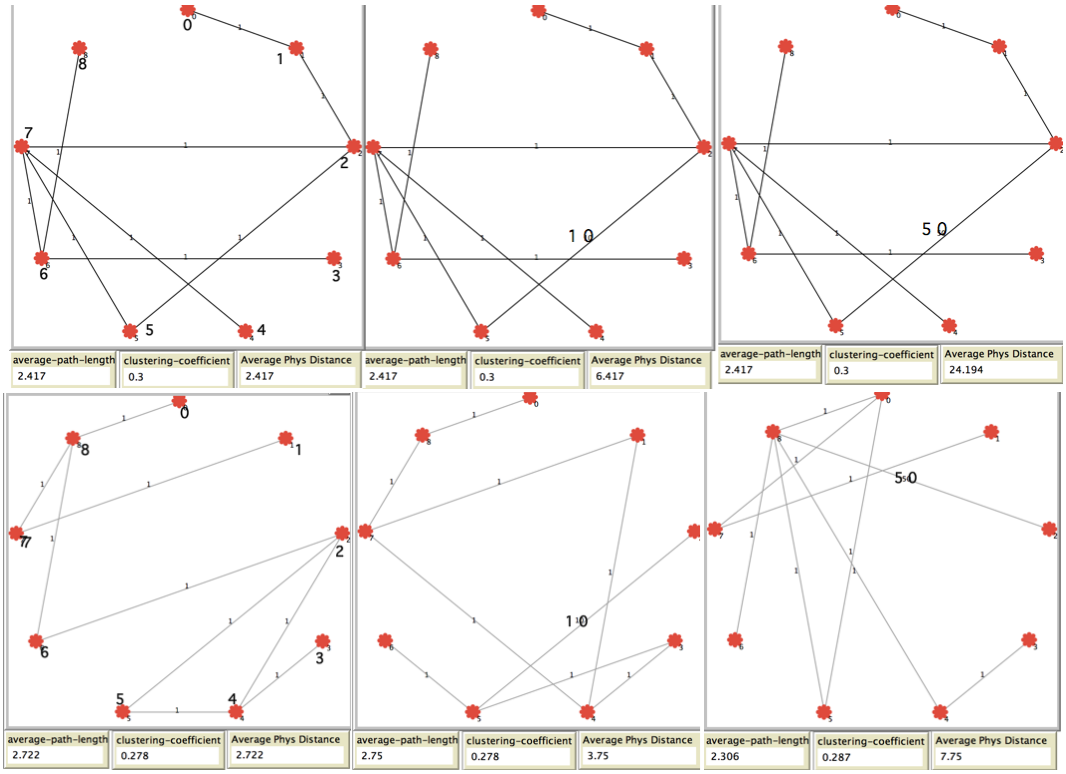


Figure 6: **A small network with 9 nodes and 9 links.**
Top panels: The Left, Middle and Right panels differ for the strength values of the link between nodes 3 and 6 (1 ,10 and 50, respectively);
Bottom panels: In the Left, Middle and Right panels the link connections were randomly redefined.

ACKNOWLEDGEMENTS

Among the many people who provided valuable suggestions in order to make the present report easier to read, I cannot miss my former PhD students Alessandro Viganó, Daiana Simone and Lisa Beatrice Caruso who were of great help in the study of neural pathologies (AV), neural networks (DS) and network algorithms (LBC).

...

REFERENCES

http://education.mit.edu/drupal/starlogo-tng.

Accornero, N. and Capozza, M. (2009). *Coscienza Artificiale: dal riflesso al pensiero*. Aracne. 8

Barabási, A. and Oltvai, Z. (2004). Network biology: understanding the cell's functional organization. . *Nat Rev Genet*, 5:101–113. 5

Bizzarri, M., Cucina, A., Biava, P., Proietti, S., D'Anselmi, F., Dinicola, S., Pasqualato, A., and Lisi, E. (2011). Embriyonic morphogenetic field induces phenotypic reversion in cancer cells: review article. *Current Pharmaceutical Biotechnology*, 12. 2

Colosimo, A. (2008). Biological simulations by autonomous agents: two examples using the netlogo environment. *Biophysics and Bioengineering Letters*, 1(3):40 – 50. 8

Colosimo, A. (2011). Multi-agent simulators: Flexible tools to reproduce collective behaviors. *Biophysics and Bioengineering Letters*, 4(2):34 – 42. 2

De Jong, H. (2002). Modeling and simulation of genetic regulatory systems: a literature review. *J. Comput. Biol.*, 9:69–105. 8

DeToni, F. and Bernardi, E. (2009). *Il pianeta degli agenti*. UTET - Torino. 2

Dzitac, I. and Barbat, B. (2009). Artificial intelligence + distributed systems = agents. *Int. J. of Computers, Communications and Control*, 1:15–26. 2

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61. 8

Ferber, J. (1999). *Multi-Agent Systems: an introduction to Distributed Artificial Intelligence*. Addison-Wesley, New York. 2

Hegsellman, R. and Flache, A. (1998). Understanding complex social dynamics: a plea for cellular automata based modelling,. *Journal of Artificial Societies and social simulations*, 1 (3). 8

Holland, J. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to Biology, Control and Artificial Intelligence*. The MIT Press - Cambridge MA. 2

Holland, J. (1996). *Hidden order: how adaptation builds complexity*. Addison-Wesley, Reading. 2

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-words networks. *Phys. Rev. Letters*, 87:1–5. 7

Milgram, S. (1967). The small world problem. *Psychology Today*, 2:60–67. 5

Mirollo, R. and Strogatz, S. (1990). Synchronization of pulsed-coupled biological oscillators. *SIAM Journal of Applied Mathematics*, 50(6):1645–1662.

Russell, S. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice-Hall. 2

Strogatz, S., Mirollo, R., and Matthews, P. (1992). Coupled nonlinear oscillators below the synchronization threshold: relaxation by the generalized landau damping. *Physical Review Letters*, 68(18):2730–2733.

Tu, B. P. and McKnight, P. (2006). Metabolic cycles as an underlying basis of biological oscillations. *Nature*, 7:697–701.

Watts, D. and Strogatz, S. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393:440–442.

Wilenski, U. (1999). http://ccl.northwestern.edu/netlogo/. 2

Wolfram, S. (2002). *A new kind of Science*. Wolfram Media, Inc. 8

...

SUPPLEMENTARY MATERIALS.

Appendix 1.The NETLOGO code in the simulation of cell phenotype changes.

```
; ============================== SET  UP ========================================
; ===============================================================================

to setup

  ca
  set switch 0
reset-ticks
clear-turtles
clear-all-plots
  crt InitialYellowCells  [set color Yellow set shape "circle"  set size 2]      ; create a given nr of  Yellow
  crt InitialBlueCells [set color Blue set shape "x"  set size 2]                ; create a given nr of  Blue c
  set TotalCellsNr InitialYellowCells + InitialBlueCells

  setup-plot      setup-plot1     setup-plot2                                    ; set the histograms' settings
  do-plotting
                                                                    ; so we have the initial histograms
end

; ============================     GO  !!  ===================================
; ===============================================================================

to go
  tick
  if ticks mod 5 = 0    [ do-plotting ]

  ; ========================  Reckoning  KINETIC CONSTANTS ======================

     ask patches [ if White_Background? [set pcolor white] draw-field-shapes]

       ask turtles [
      if (color = Yellow)  and (random-float 1000 < (Yellow_to_Blue-k_intrinsic + InRadiusYellow * YELLOWCrowdC
          set color Blue set shape "x"  set size 3                                                              ]

      if (color = Blue)  and (random-float 1000 < (Blue_to_Yellow-k_intrinsic + InRadiusBlue * BLUECrowdCoeffic
          set color Yellow set shape "circle"  set size 2                                                       ]

  ; (( CROWDING effect ))

    if (color = Blue)  [set InRadiusBlue count turtles with [color = Blue] in-radius RADIUSsize]

    if (color = Yellow)  [set InRadiusYellow count turtles with [color = Yellow] in-radius RADIUSsize]

  ; ===============================================================================

                ]

  ; ===============================================================================


   ask patches  [  draw-field-shapes]

   ask turtles [
    if color = Yellow  [ rt random 360 sense-field fd 2 if (distancexy 0 0 > max-pxcor / 2)  [back 2]   ]
    if color = Blue [ rt random 360  sense-field fd 2 if (distancexy 0 0 > max-pxcor / 2)  [back 2]     ]

end
```

...

Appendix 2. Definition of the network parameters used in this work.

Name	**Symbol, Definition**
Node Degree	K_i = number of links of node i (*)
Shortest Path Length (**)	$SPL_{i,j}$ = minimum number of links going from node i to node j: APL = *Average Path Length* = $< SPL >$ over all pairs of nodes
Clustering Coefficient (***)	$cc_i = \frac{2e_i}{K_i(K_i-1)}$, where e_i = connected pairs within the first neighbors of node i $< cc >$ = average over all nodes having K > 1. (see also Eq. 4).

(*) In a network without self-loops (connecting a node to itself) and multiple links (two nodes connected by more than one edge), K_i also equals the number of neighbors of node i. In networks where each link has a selected direction there is an incoming degree, K_{in}, which denotes the number of links that point to a node, and an outgoing degree, K_{out}, which denotes the number of links that start from it.
(**) The distance (path) between two nodes is measured by how many links must be passed through to travel between such nodes. The shortest path, the path with the smallest number of links between the selected nodes, has a special role.
(***) In many networks, if node A is connected to B and C, then it is highly probable that B also has a direct link to C.

GOOD DESIGN PRACTICE FOR MEDICAL DEVICES AND EQUIPMENT

A COMSOL MULTIPHYSICS Approach To 'Design For Validation'

Ole Brix[1], Huifang Zhao[1], Shiming Yang[1], Eivind Olav Dahl[1], Louise Mohn[1], Kjetil Folgerø[2], and Jan Kocbach[2]

1. *Luzmon Medical AS, Bergen, Norway*
2. *Christian Michelsen Research AS, Bergen Norway*

Keywords: ComsolMultiphysics, electric muscle stimulation, medical regulations, EU Counsel directive 93/42, FDA

Abstract: In the European Union (EU) and the United States (US) regulation of medical deviceshasevolved significantly the last 25 years. The attention has moved from focusing only on manufacturing processes towards device design and manufacturing processes. As a result the medical device industry has been forced to take a m ore concurrent and integrated approach to design control, a k ey element internationally in medical device regulations, to focus on design, development and validation. Unfortunately these regulations rarely provide any useful guidance on howto produce the evidence to meet the new medical device requirements. In this respect current good design practice (CGDP) needs to be extended to include 'Design for Validation' to enforce designers to consider validation during the design of both the device and its related process equipment.

Luzmon Medical AS has with its research and development partners used COMSOL Multiphysics(2014) in the designing process of new electrodes and a device for muscle stimulation intended for rehabilitation. The software was used to simulate the heat distribution and electric field distribution within body tissue for varying design parameters and body compositions and in this respect validate the design.

The current study thus presents the results of simulating a physiological system within this framework.

1 Introduction

A medical device is inthe EU Council directive 93/42 (European Council Directive 93/42/EEC) defined as: 'Any instrument, apparatus, appliance, material, or other article, whether used alone or in combination, including the software necessary for its proper application, intended by the manufacturer to be used for human beings for the purpose of:

- Diagnosing, prevention, monitoring, treatment or alleviation of disease.
- Diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap.
- Investigation, replacement or modification of the anatomy or of a physiological process.
- Control of conception,

and, which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted in its function by such means.'

The design and development have been strongly regulated by medical directives and international standards since 1944.We find minor differences in the objectives of the EU and US regulations..Where the objective of the EU regulation is to ensure that the devices are safe, and that they perform as the manufacturer intended, the objective of FDA regulation (USA) is to ensure that the produceddevices will be safe, effective, and in this respectpresent a b enefit to the user. Both regulations points out that product development is inherentlyan evolutionary process, in which validation will ensure that a d evice is fit for purpose, and that safety and effectiveness in this respect cannot be proven by final inspection or testing.The quality management system requirements are inthe EU regulations met through ISO 9001and ISO 13485, and theUS requirements are met through the Quality System Regulation (QSR), which is part of the Code of Federal Regulations. In the context of QSR, ISO 9001 and ISO 13485,validation thus ensures fitness for purpose by providing documentary evidence that the device, premises, plant, equipment, process and test methods are capable of functioning, and continuing to function,in accordance to their design and specification. In this respect validation is concerned with demonstrating the consistency and completeness of a design with respect to the user requirements, the initial ideas that outlines the purpose and function of the device.Validation should not be confused with the verification process thatis meant to ensure thatthe output from each phase in design and development meetsthe input requirements specified in the output of the previous phase.

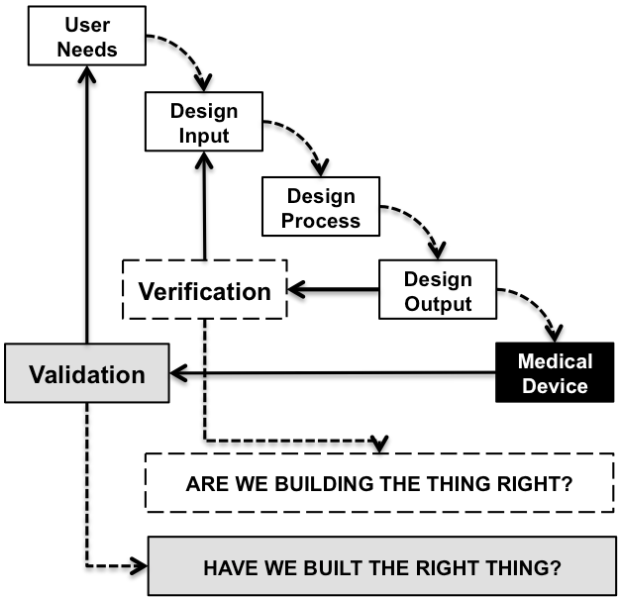


Figure 1: The waterfall model (GMP, 1976).

The concept of validation has been introduced in literature in the form of the Waterfall Model (Fig. 1), which describes a sequential design process, in which progress is seen as flowing steadily downwards (like a waterfall) through the phases of conception, initiation, analysis, design, construction, testing, production/implementation, and maintenance.

The Production Development[1] Validation V-Model Fig. 2) was designed to produce a simplified understanding of the complexity associated with systems development to detailed, rigorous development lifecycle models and project management models.Each of these models validation is clearly outlined, but only in relationship to specific aspects in the medical device development.In order to clarify some of the concepts found in the official definitions, FDA included the Waterfall Model in its document entitled Design Control Guidance for Medical Device Manufacturers (GMP, 1976).

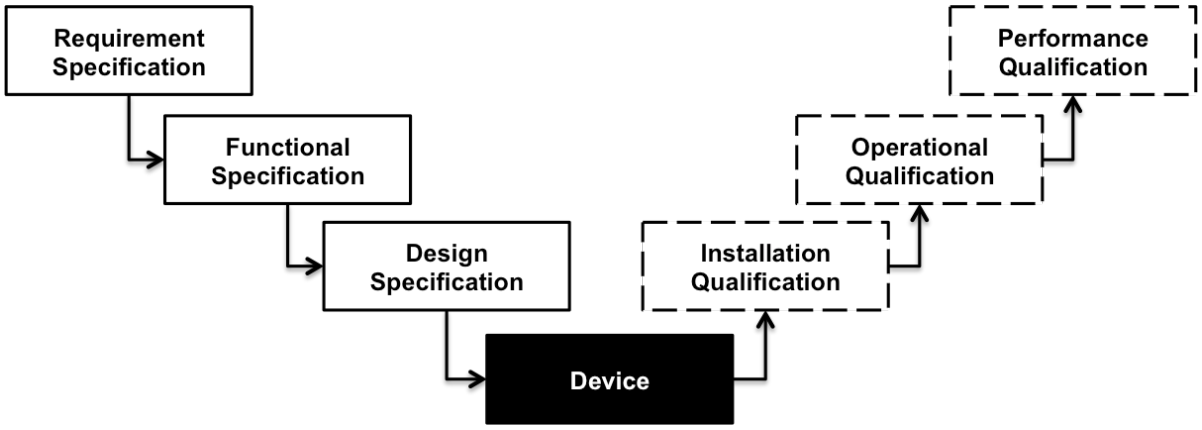


Figure 2: The Production Development Validation V-Model.IQ[2], OQ[3] and PQ[4] relate to three levels of specifications (see footer).

The waterfall model indicates that verification establishes that the design output meets the requirements encapsulated in design input. Verification is a detailed examination of aspects of a design at various stages of development, while design validation is a cumulative summation of all efforts, including design verification, and extends to the assessment to safeguard that devices manufactured in accordance with the design actually satisfy user needs and intended use. The Waterfall Model thus does not show validation as it applies to the process of equipment design and development

The Production Development Validation V-Model outlines validation as it relatesto the development of the production equipment in the form of, Installation Qualification (IQ)[5], Operational Qualification (OQ)[6], and Performance Qualification (PQ)[7],(Fig. 6). In this respectit provides a picture of the stages missing in the

[1] The development of the process equipment or procedures used tomanufacture, assemble or test a medical device.
[2] Design Specification is a complete definition of the equipment or system in sufficient detail to enable it to be built. This links to Installation Qualification, which checks that the correct equipment or system is supplied; that it meets the required standards and that it is installed correctly.
[3] Functional Specification is documentation, normally written by the supplier describing the detailed functions of the equipment or system (i.e. what the system will do). This links to Operational Qualification that tests all the functions specified.
[4] Requirements Specification describes what the equipment or system is supposed to do and, as such, is normally written by the pharmaceutical manufacturer. This links to Performance Qualification that tests these Requirements.
[5] Documentation demonstrating that the equipment designs and configuration is as intended, that instrumentation has adequate accuracy, precision and range for intended use and that services (such as power supplies) are of adequate quality.
[6] Documentation demonstrating that the equipment or system operates as intended throughout representative or anticipated operating ranges.
[7] Documentation demonstrating that when operated within set parameters the process will consistently produce a product meeting its predetermined specifications.

Waterfall Diagram, the Production Development Validation V-Model does however not represent validation or verification as it applies to design. Since the specifications of the V-Model are outputs of the design process the qualification activities are shown to occur after the device and process equipment have been designed and built.

1.1 Design for Validation Model Derivation

In (Figure 3 we have outlined (grey boxes) the major steps in transforming user needs to the final medical device on the left hand side of the diagram in terms of device design, process design and production development.These steps do often occur in parallel, but arein the figure outlined serially for simplicity. The development of validation requirements should commence as soon as the user needs are known.Device validation encompasses all activities within the large "Device Validation V" on the diagram - device design, process design, and production development.Process validation involves all of the activities within the smaller "Process Validation V"diagram part including - process design and production development.

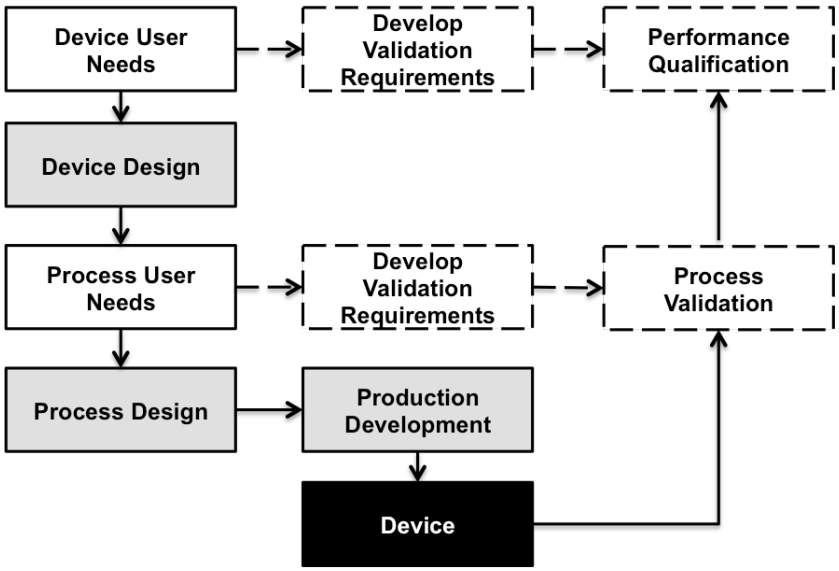


Figure 3: Validation as it applies tothe Design and Development Process (based on Clarkson, Bishop, and Fox, 2001).

The prototype manufacturing equipment (device) used for verification during process design ultimately becomes the final manufacturing equipment at the end of production development.

The device is finally validated against user needs. Process user needs consist of the device design specification and the process equipment requirements,which are often derived during the device design process.Device user needsare not so clearly defined. They must be derived from a multitude of different sources. In this respect device validation can be much more complicated than processvalidation.Validation performed after production and verification of the final device may involve clinical trials or evaluations to ensure that the final product meets the original user needs and intended uses depending on compliancy with predicative devices.

The verification process occurs within each of the device design, process design and production development activities in order to ensure that'We are we building the thing right?'

The verification process is outlined in more detailsin Figure 4.In this process we willprove that design outputs meet design inputs.The figure shows the verification process as a sequentially process during the iterations of the design and development activities. Verification requirements can as validation requirements be developed in parallel with design inputs development.The results of design verification are used for'final device design verification', and the design output is used to confirm 'process user needs'.

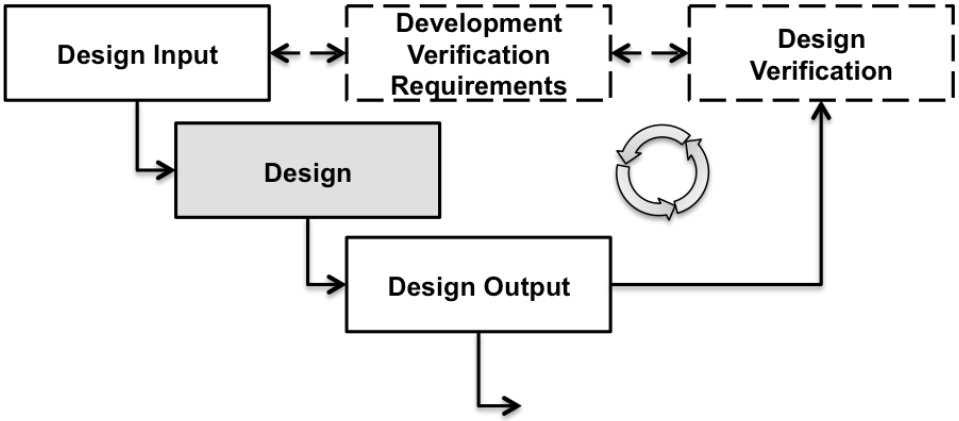


Figure 4: Diagram of verification requirements(based on Clarkson, Bishop, and Fox, 2001).

Process qualification is a means of proving that process development output meets process development inputduring the development of all the manufacturing and test equipment. IQ, OQ, and PQ (Figure 2)are related to the Production Development Validation V-Model shown previously.The process qualification requirements and the results of process qualification provide important input to the 'final process design verification'. When the production processes has been qualifiedit can be used to manufacture the final device.

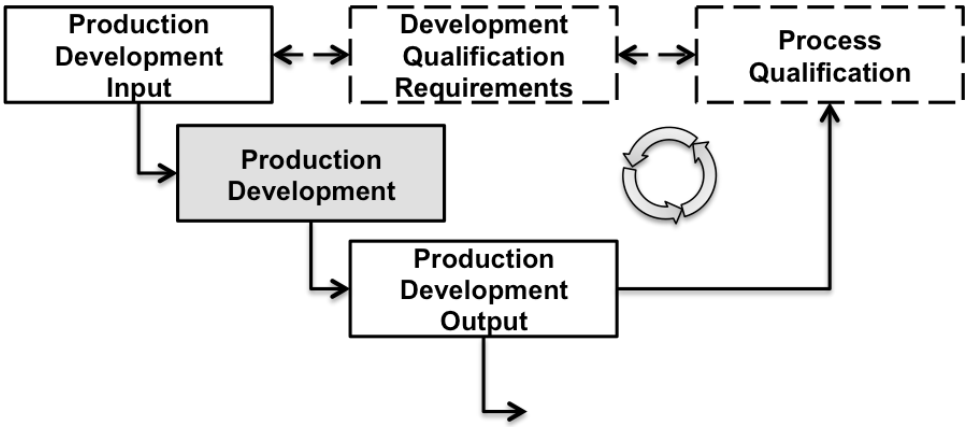


Figure 5: Diagram of process qualification(based on Clarkson, Bishop, and Fox, 2001).

The DFV V-Model clearly shows that qualification of the process equipment is a subset of both process validation and device validation.

Luzmon Medical AS has implemented the design controls as design reviews in their 'Risk Management Process' as outlined in Figure 6and

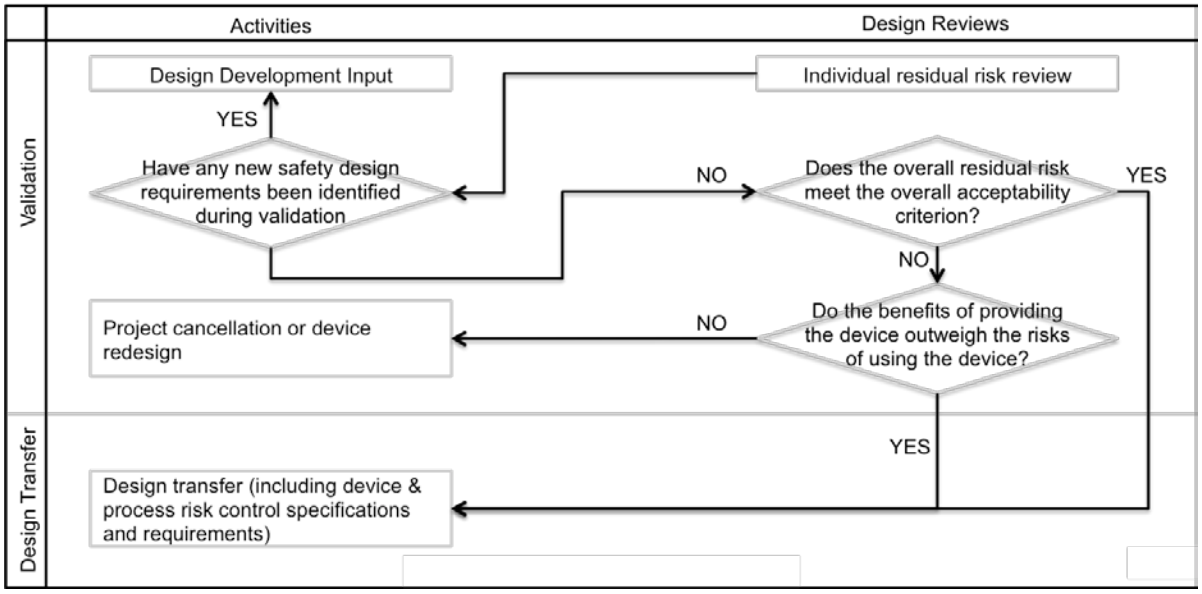


Figure 7and in compliance with both the Design for validation V-Model and the Risk Management standard ISO 14971.

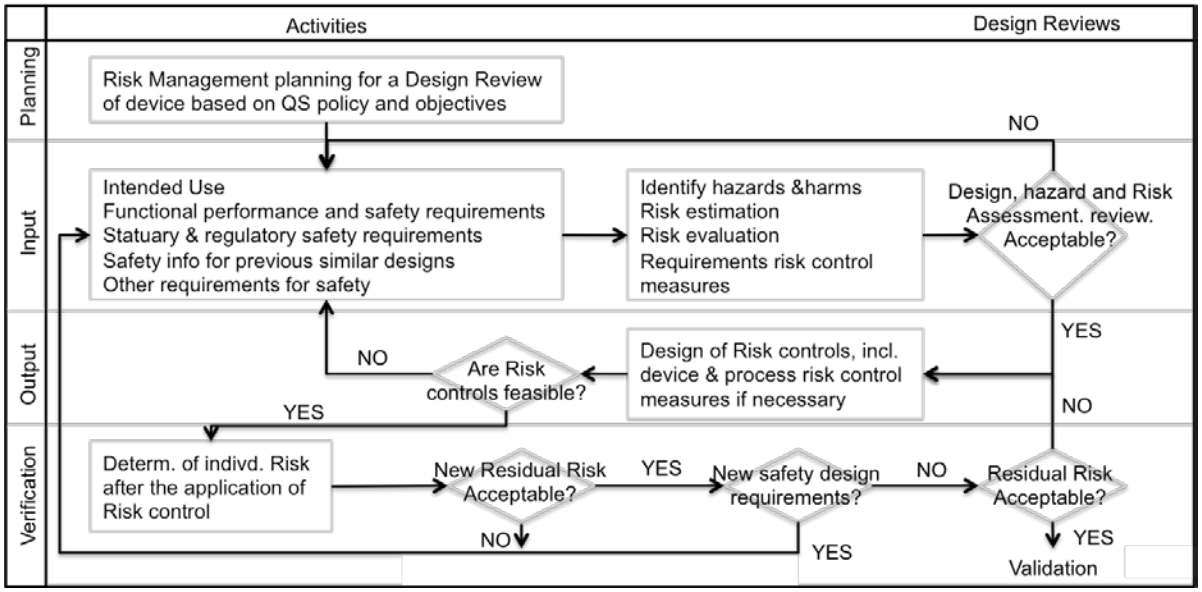


Figure6: The Design and Development process diagram (outer shell) for LuzmonMedical AS showing the following four Life Cycles: 1) Planning, 2) Input, 3) Output, 4 Verification.

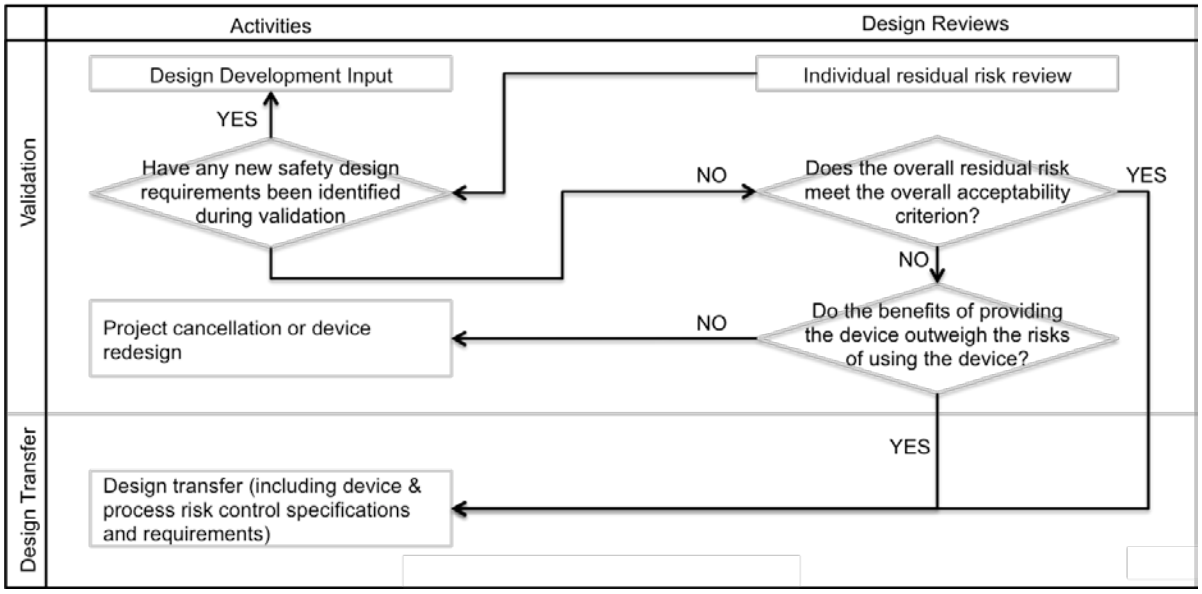


Figure 7: The Design and Development process diagram (outer shell) for LuzmonMedical AS showing two further Life Cycles: 1) Validation, 2) Design Transfer. Beneath the outer shell diagram we consider for a medical device development in more details validation of i.e. bio-compability, usability, reprocessing and other critical process validation, software.

2 Use of COMSOL Multiphysics for validating design

The finite element simulation tool COMSOL Multiphysics has been used in order to validate different aspects of the design. COMSOL Multiphysics is widely used and accepted for verification and validation of design (see their website: www.comsol.com). This is important to medical device designers working with experimental design activities involving the use of animals, which is regulated by the European Directive 86/609/EEC.

In our study we have simulated heat distribution within body tissue in order to validate the significance of a "user property", fat layer thickness, for the absorbed heat and heat distribution. The technology for applied heating was radiation (traditional technology) using heat wire separation (this study), infrared technology (IR;Kocbachet al., 2011), and microwave (Kocbachet al., 2011). We have carried out electric muscle stimulation simulations in order to validate the conductivityof thermoplastics (TPU; Kocbachet al., 2011), and the possible impact of using conductive gel as an interface layer between skins and electrodes. We have used the AC/DC module of COMSOL Multiphysics for the electromagnetic (E bio-compabilityM) simulations and the Heat Transfer module for the heat simulations.

2.1 Model description

The body tissue in the COMSOL models has been modelled as a layered structure, consisting of a bone layer, a muscle layer, a fat layer and a skin layer. Electrodes, heat pads and conductive gel have been modelled as separate layers on top of the skin layer. The heating and heat transfer simulations are mainly done using an axisymmetric model whereas the electrode stimulation models are made using a 3D model. COMSOL LiveLink for Matlab has been used in order to do script based parametric analysis, sensitivity analysis and iterative optimization from Matlab.

2.1.1 Heating and heat transfer simulations

The heating and heat transfer simulations are done using the Multiphysics capabilities of COMSOL Multiphysics where the AC/DC module and the heat transfer module is coupled. The AC/DC module is applied for simulating the effect of generating heat from heat wires and Penne's Bio-heat equation in the heat transfer module is used for heat transfer into the body. An iterative two-step process is applied in order to find the appropriate electric potential required to heat the body tissue to a surface temperature of 42 ºC at the parts of

the heat pad, which are in contact with the skin. First electromagnetic simulations are made in which the electric potential distribution, electric field distribution and electromagnetic losses in the heat pad are calculated. This is followed by a heat simulation in which the heat distribution in the tissue and the heat pad is calculated based on the heat generated by the electromagnetic losses in the heat wires. The electric potential is changed until the correct surface temperature is obtained.

Penne's Bio-heat equation (Pennes 1948) is given by

$$\rho C_p \frac{\partial T}{\partial t} + \nabla \cdot \left(-k \nabla T\right) = \rho_b C_b \omega_b \left(T_b - T\right) + Q_{met} \tag{1}$$

where

ρ =tissue density [kg/m^3]
C_p=tissue heat capacity [J/kgK]
T=tissue temperature [K]
k=tissue thermal conductivity [W/mK]
ρ_b=density of blood [kg/m3]
C_b=heat capacity of blood [J/kgK]
ω_b=blood perfusion rate [1/s]
T_b=arterial blood temperature [K]
Q_{met}=metabolic heat generation [W/m^3]

The bioheat equation consists of 4 terms representing heat storage, heat conduction, heat convection (blood perfusion) and metabolism. The metabolic heat generation is the sum of the basal metabolic rate and local autonomic thermoregulation (exercising and shivering) (Fialaet al. 1999).

There are some limitations related to the fundamental assumptions of the Pennes model (Kreith&Goswami 2005), and several extensions and modifications of the Pennes model have been suggested. For example, Bronzino (2000) claims that in order to get realistic results, $\omega_b$and T_b should be taken as adjustable parameters determined by e.g. curve fitting, and not literally as blood perfusion rate and arterial blood temperature. However, the model has nevertheless proved to be in good agreement with experimental findings and is the primary choice for heat transport analysis (Kreith&Goswami, 2005), and thus the Pennes model is used for the simulations in the current work.

The heat wires are modelled using the AC/DC module of COMSOL Multiphysics.Heat wires have been simulated by varying the electrical conductivity in the heat wire layer using a step function. An electric potential was applied over the heat wire layer (+V_{pad} at the upper end of the model, ground at the lower end of the model). This resulted in heat being generated in the parts of the heat wire layer for which the electrical resistance is high, i.e. in the heat wires. The electrical conductivity in the heat wire layer was varied according to the formula

$$y = \sigma_{resistive} + \frac{\sigma_{metal}}{2}\left\{sgn\left[\cos\left(2\pi N \frac{z}{l_{model}}\right)\right] + 1\right\} \tag{2}$$

where l_{model} is the length of the model, $N$is the number of heat wires to be applied along the length of the model, σ_{metal} is the electrical conductivity of the part of the heat pad in which no heat is to be generated (heat wire spacing) and $\sigma_{resistive}$is the electrical conductivity of the part of the heat wire layer in which heat is to be generated (heat wires).

The simulation model used for the heating and heat transfer simulations is an axisymmetric model (Figure 8) built up using uniform layers. The layers as seen from the centre of the model are bone, muscle, fat and skin (too thin to be properly identified in the axisymmetric model). Outside the skin layer there is a heat pad structure consisting of a silicon rubber layer and a heat wire layer. The model has been built up to represent a

part of a typical thigh in the human body, but results are also applicable to other body parts. The length of the model was set to 200 mm, the radius of the bone to 20 mm, muscle layer thickness to 50 mm, the fat layer thickness was varied between 1 mm and 100 mm (nominal value is 10 mm) and the skin layer thickness was set to 1 mm. The thickness of the silicon rubber layer was set to 2 mm and the thickness of the thin heat wire layer was set to 0.5 mm.

One important issue when setting up the simulation model has been to find appropriate material parameters, especially for biological tissue. Biological tissue properties found in the literature vary significantly between different sources. Also, it must be noted that the blood perfusion rate and metabolic heat generation depends strongly on the metabolic and thermoregulatory needs of the tissue, and the bio-heat equation has fundamental limitations. In addition the tissue layers are assumed to be isotropic, i.e. the anisotropy in the tissue is not included. Kuhn & Keller (2005) showed that the significant differences in tissue parameters for muscles in the longitudinal direction compared to the transversal direction have a notable influence on the simulation results. Thus the focus for the simulations has been to get a qualitative understanding of the processes involved, and the outcome of the simulations must therefore be interpreted with great care. Heat pads have been modeled as additional outer layers of silicone rubber and heat wire layer (Figure 8).

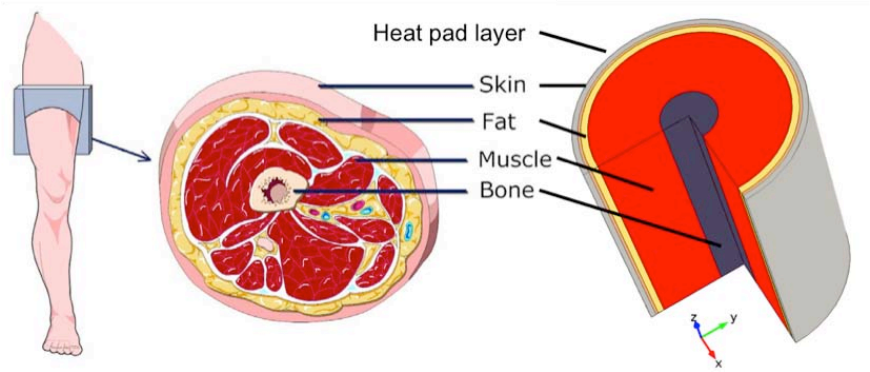


Figure 8: Heating and heat transfer model.

Tissue parameters were set according to the parameters defined in Table 1.

Table 1: Properties of biological tissues used in the present work from *Carrara (2010), †Duck (1990), ‡Drizdal et al. (2010) and §Fiala et al. (1999).

		Blood	Bone	Fat	Muscle	Skin
Thermal conductivity†	[W/mK]	0.492	0.35	0.25	0.5	0.32
Spesific Heat†	[kJ/kg°C]	3.8	1.3	3.1	3.72	3.45
Density†	[kg/m3]	1050	1990	928	1041	1100
Blood perfusion rate§	[m3/s/m3]	-	0	$0.0036 \cdot 10^{-3}$	$0.538 \cdot 10^{-3}$	$1.05 \cdot 10^{-3}$
Basal metabolic heat rate§	[W/m3]	-	0	58	684	368
Conductivity* 10 kHz	[S/m]		0.020	0.024	0.341	0.003
Conductivity 434 MHz	[S/m]		0.022†	0.042‡	0.8‡	0.68‡
Rel. Permittivity* 10 kHz			$0.5 \cdot 10^3$	10^3	$25 \cdot 10^3$	$29 \cdot 10^3$
Rel permittivity 434MHz			90†	5.6‡	57‡	49‡
Optical attenuation coefficient**				340 (λ=630 mm)	370 (λ=800 mm)	1000 (λ=1230 mm)

Properties for the pad materials are shown in Table 2. Temperature dependency of tissue parameters has not been taken into account in the simulations.

Table 2: Properties of pad materials used in the present work from [*]Watlow (2010), [†]Primasil (2010), [‡]Matweb (2010) and [§]Ward (1984).

Material	Conductivity [S/m]	Rel. Permittivity	Thermal conductivity (W/mK)	Spesific Heat (kJ/kg°C)	Density (kg/m3)
Stimulation, electrodes (conducting silicon rubber)	40[†]	7[†]	0.216[*]	1.884[*]	1250[*]
Stimulation, isolating layer .	1.6e-15[‡]	7[‡]	0.216[*]	1.884[*]	1250[*]
Heat pad layer	1.6e-15[†]	7[†]	0.216[*]	1.884[*]	1250[*]
Conducting Gel (various).	0.11 - 6.2[§]	(80)[§]			

2.1.2 Electrical stimulation simulations

The simulation model for electrical stimulation is also a 3D layered model (Figure 9), but in order to include the stimulation electrodes the model is not axisymmetric as was the case for the heating simulations. From the bottom and up the layers are given as a 10 mm bone layer, a 30 mm muscle layer, a 5 mm fat layer, a 1.5 mm skin layer and a 1 mm stimulation pad layer. In some of the simulations a 1 mm gel layer is included between the skin layer and the stimulation pad layer. The stimulation layer contains two excitation electrodes. For the given case, the electrode length is 190 mm, electrode width is 50 mm and electrode spacing is 50 mm. A strip between the electrodes is used for gel leakage simulations. The electrodes and heat pads were modeled as separate thin layers on top of skin layer.

Four different model variants were considered; a model with finite thickness electrodes with varying electrical conductivity, a model with thin high-conductivity electrode wires below the electrodes, a model with a conductive gel layer with “gel leakage” between the electrodes and a model with non-conductive air-layer between the electrodes.

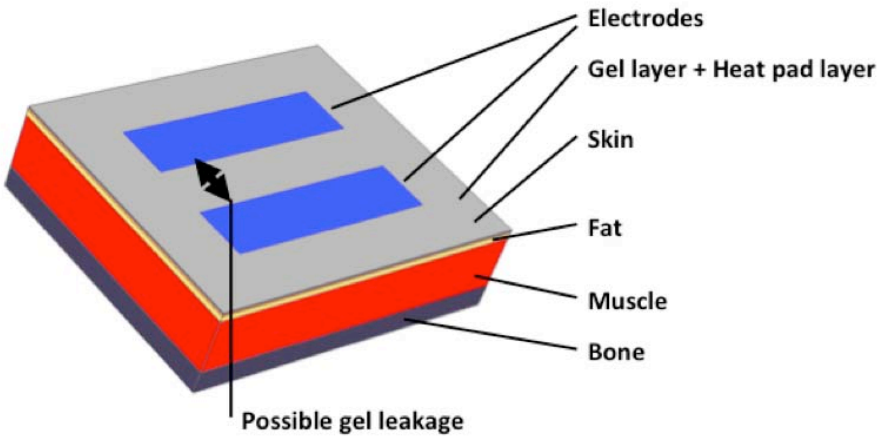


Figure 9: Electrical stimulation model Electric field applied at short edge of electrodes (either on same side or opposite sides).

2.2 Results and discussion

2.2.1 Validation of heating related to 'user properties' - fat layer thickness

Usability engineering(Figure 10) is an important part of the design and development process of medical devices,closely integrated with the risk management process as outlined in Figure 11.

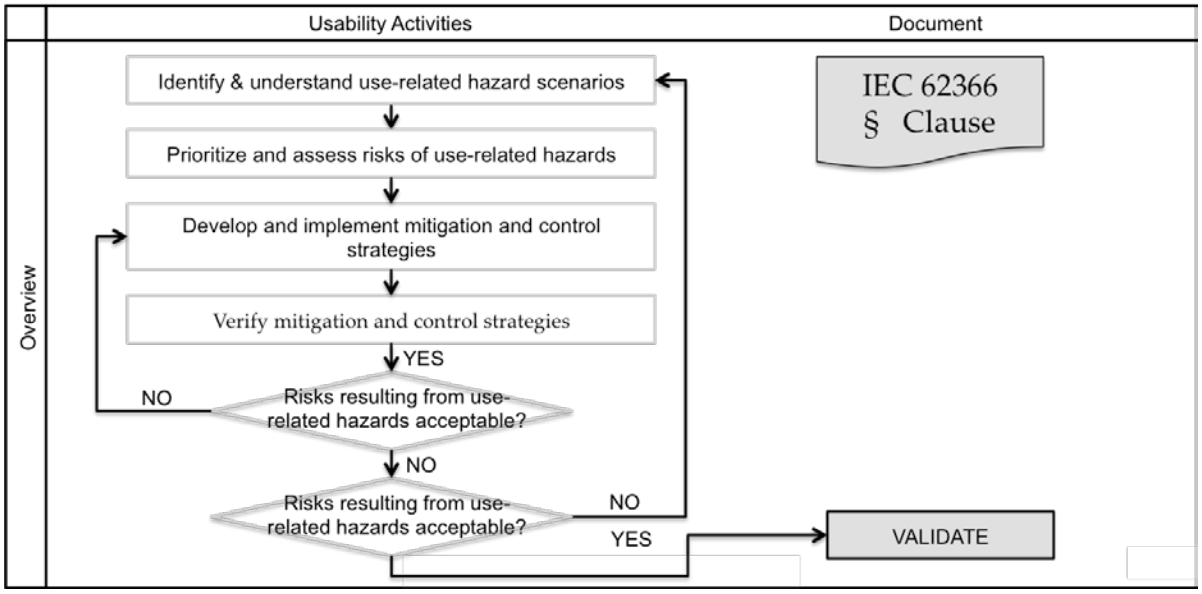


Figure 10: Usability engineering diagram - Human factors (HF).

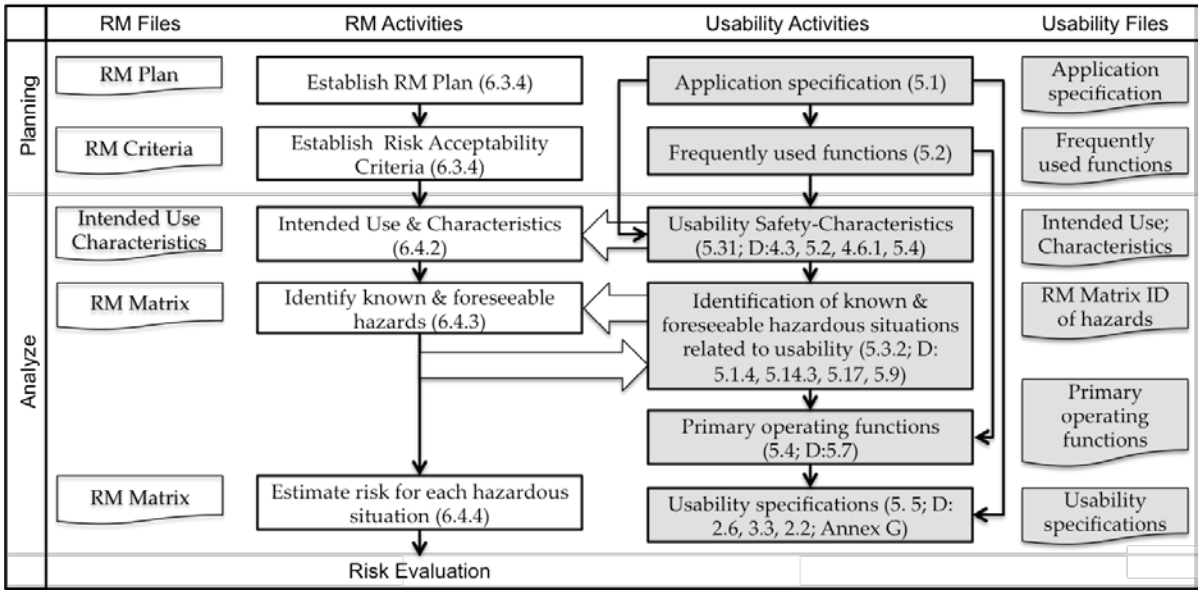


Figure 11: Risk analysis diagram outlining the integration of usability with design and development and risk assessment. High-level illustration only; in practicethe combined risk management and usability management process is far more detailed and encompass more than 40 interrelated activities.

The intended use is the most important usability input setting the framework for regulatory requirements and thus design and development, risk management, post market surveillance, etc.

The intended use of amedical device could for example be adjunctive therapy in rehabilitation for medical purposes. Based upon this intended use, the medical device manufacturer should specify indications where the medical device in question can successfully be used, for example the six most common indications for electrical muscle stimulators approved by FDA (Food and Drug Administration, US):Relaxation of muscle spasms, prevention or retardation of disuse atrophy, increasing local blood circulation, muscle re-education, maintaining or increasing range of motion, and immediate post-surgical stimulation of calf muscles to prevent venous thrombosis.

Generally, electrical stimulation (EMS) is employed to provide exercise patterns when patients are unable to perform such exercise voluntarily due to: Pain, restriction of joint motion, dysfunction of the neuromuscular system, and weakness or disuse atrophy.

This implies that we have to consider a patient group with large differences in body tissue composition when simulating temperature therapy as well as EMS.In this respect we have focused on the significance of body fat composition for validating heating technologies with respect to efficiency and penetration into the body tissues. Previously we have concluded that there are no differences in using direct heating from resistance wires or IR, and that deep tissue penetration of temperature can only be achieved by using microwaves, which has been out of the range of the current development, but could play a significant role in i.e. treatment of cancer (Kobach et al., 2011).

In this study we have therefore focused on analyzing the possible impact of the thickness of the fat layer below the skin to validate the required period for pre-heating muscles before stimulation as well as the of effect of temperature penetration.

The simulations show significant variation of both heating time and final temperature with fat layer thickness, and thus it is important to adjust the time needed for preheating of a patient before starting muscle stimulation according to the patient's body composition.

The time it takes to heat up the muscle tissue increases when going into the muscle. Figure 12 shows the variation in heating time at certain points within the muscle and on the muscle/fat boundaryfor a fat layer of 10 mm. After 20-30 minutes the temperature is around 0.5 ºC below the steady-state temperature. At the muscle/fat layer interface 72% of the temperature increase happens within the first 20 minutes and 84% of the temperature increase happens within the first 30 minutes. Considering a position 20 mm into the muscle, only 30% of the temperature increase happens within the first 20 minutes. To get 80% of the temperature increase, a heating time of 50 minutes is required. This illustrates that a longer heating time is required to obtain deep heating of the muscles.

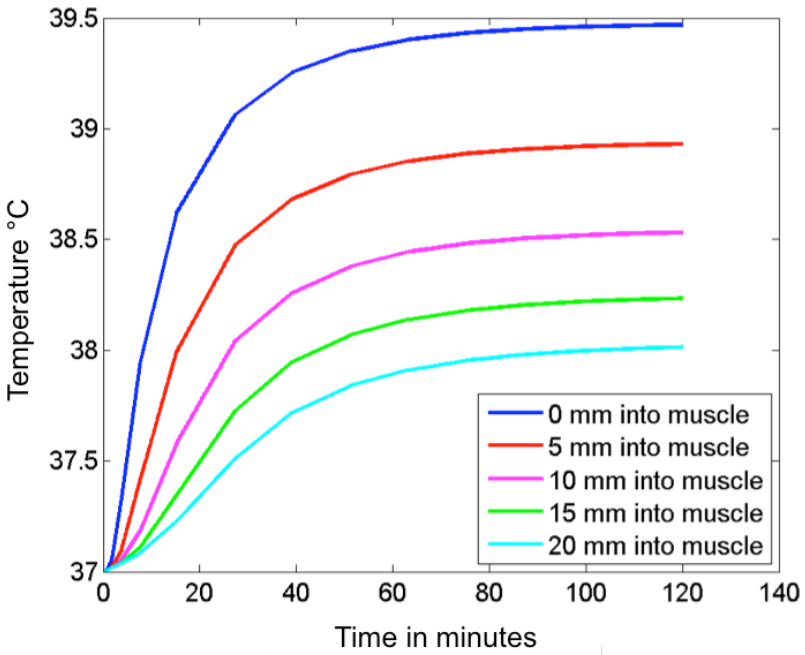


Figure 12: Temperature variation as a function of time at muscle/fat layer interface as a function of time for fat thicknesses of 1- 40 mm. It should be noticed that only the basal metabolism is included in these simulations, which implies that the time for heating the muscles would be faster if stimulation could start earlier and thereby increase metabolism.

Heating pads normally contain a long resistance wire that is curled up to cover the most of the pad area evenly with a fixed distance between the parallel running parts of the wire. In this respect wire thickness (and resistance characteristics) as well as distance between the parallel running parts are key parameters to adjust regarding optimizing heat distribution.

To validate these features we have used a functional model where the surface area is divided into equally sized regions with equal spacing, for which heat is applied to achieve a temperature of 42 ºC at the outer edge of the skin. For all the simulations we have used a fat layer thickness of 10 mm, and the blood temperature is assumed to be constant (37 ºC). Furthermore the effect of transportation of blood heated by the wires to regions without heat wires is therefore not included. In the current simulation series, the heat wire spacing is set equal to the heat wire width. If the heat wire width is set larger than the heat wire spacing, a more even temperature distribution can be achieved with larger heat wire spacing.Figure 13 show the case with two 5 mm wide heat wires separated by 5 mm spacing (left), and with eight 1.25 mm wide heat wires separated by 1.25 mm spacing (right).

Heat wires have been simulated by varying the electrical conductivity in the heat wire layer using a step function. An electric potential was applied over the heat wire layer ($+V_{pad}$ at the upper end of the model, ground at the lower end of the model). This resulted in heat being generated in the parts of the heat wire layer for which the electrical resistance is high, i.e. in the heat wires. The electric potential $V_{pad}$was adjusted using an iterative process in order to get a maximum temperature of 42 ºC at the parts of the heat pad whichare in contact with the skin. For each iteration in the iterative process, two simulations were made. First an electromagnetic simulation in which the electric potential distribution, electric field distribution and electromagnetic losses in the heat pad were calculated. This was followed by a heat simulation in which the heat distribution in the tissue and the heat pad was calculated based on the heat generated by the electromagnetic losses in the heat wires, where the electrical conductivity was varied according to eq. (2)

The heating effect of the two-wire cases is very uneven – giving a temperature variation of more than 1.2 ºC at the fat/muscle interface. The effect is lower at deeper muscle positions, but a significant reduction of the heat compared to the uniform heating case is observed. The maximum heat loss compared to the uniform case is 1.5 ºC at the fat/muscle interface and about 1.0 ºC at a position 10 mm into the muscle. For the case with eight heat wires with heat wire width of 1.25 mm and 1.25 mm heat wire spacing the maximum heat loss compared to the uniform case is only 0.4 ºC at the fat/muscle interface and 0.15 ºC at a position 10 mm into the muscle.

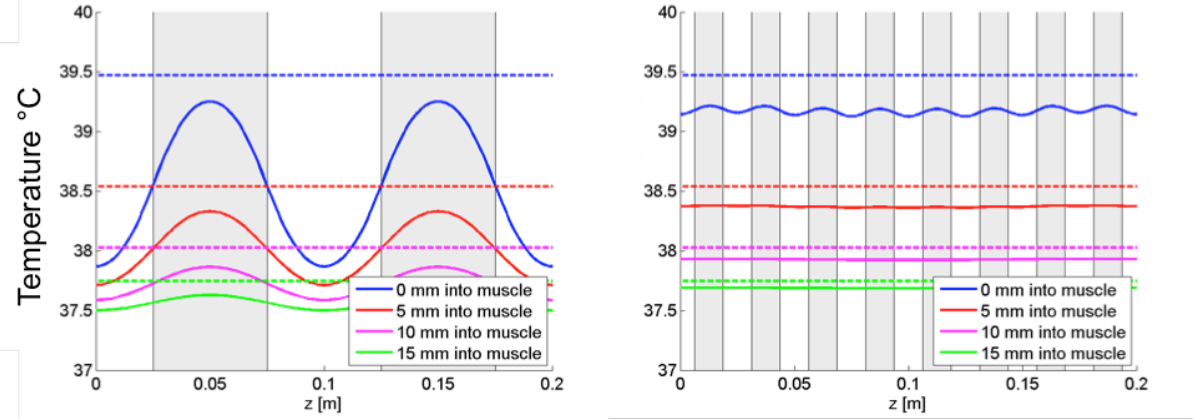


Figure 13: Heat distribution along the length of model at the fat layer/muscle layer interface and at three different positions in the muscle tissue. Solid lines are for the cases with heat wire spacing of 10 mm (left) and 1.25 mm (right). Dotted lines are for uniform electrodes.

In this respect the simulations revealed that the heating technology was adequate when optimized with respect to wire width and spacing, also considering the energy applied by stimulation, however we were not completely satisfied by the general heat distribution pattern and decided to develop a new design using the same technology, but allowing a better spatial distribution of heat. This approach has now been patented worldwide.

2.2.2 Validating the use of conductive TPU as electrodes for EMS

The conductivity of the TPU electrodes is expected to have a significant influence on the electric field and current applied to the muscle, and thus not only the quality of the treatment with respect to transducing pulses but also the well-being of the patient during the process of treatment. An uneven electric field distribution may cause under-stimulation in one end and over-stimulation in the other end of a muscle passing the pain threshold of the patient.

To avoid this unwanted design case we must assure that our design can produce a near uniform electric field and current in the muscle.This could appear to be a quite impossible task, since we want to maximize conductivity in the electrodesnot jeopardizing flexibility of the pad that actually will decrease with increased conductivity of the TPU material.When the electrode conductivity is low, the electric potential will in the general case vary significantly along the electrode, especially if the contact points for the electrodes are at the same edge of the electrodes. Our challenge was to invent a design that could counteract this not wanted scenario using a relative low conductive TPU with reasonable flexibility.

To investigate the effect of electrode conductivity, COMSOL Multiphysics simulations with varying parameters for the conductivity of the electrode were done. Simulations were done for conductivities ranging from 1 MS/m (metal) to 0.1 S/m (poor conductor). The results showed that when the conductivity decreases, the electric field is focused in a small region close to one end of the electrodeas expected (Figure 14). This gives very low values of electric field at the other end of the electrode, leading to difficulties in achieving the “critical electrical field” for stimulation in this part of the muscle while being below the pain threshold of the patient.

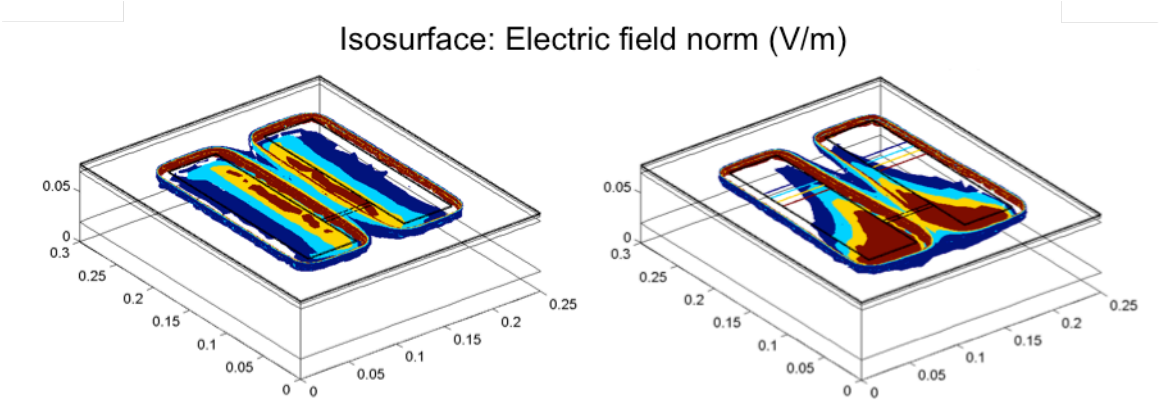


Figure14: 3D ISO-surface plot of the electric field in the tissue with varying electrode conductivity. ISO-surfaces are plotted for electric field values of 10 V/m, 12 V/m, 14 V/m and 16 V/m. Electrode conductivity of 1 MS/m (left) 10 uS/m (right).

A method to compensate for poor electrode conductivity, however, is to integrate a high conductive wire under the TPU electrodes. This scenario was simulated using the COMSOL Multiphysics. The shape of the wires was chosen in order to investigate a worst-case situation in which the distance between the wires with high conductivity is large at one end of the electrodes and small at the other end of the electrodes. This is illustrated in Figure 15. Here only two electrode conductivities are considered – the metal case with conductivity of 1 MS/m and the case with a very poor conductivity of 10 μS/m.

From the figure it is evident that when a conductive wire is used below the electrodes, there is no significant variation in stimulation along the electrode, even if very poor electrode conductivity of 10 μS/m is used.

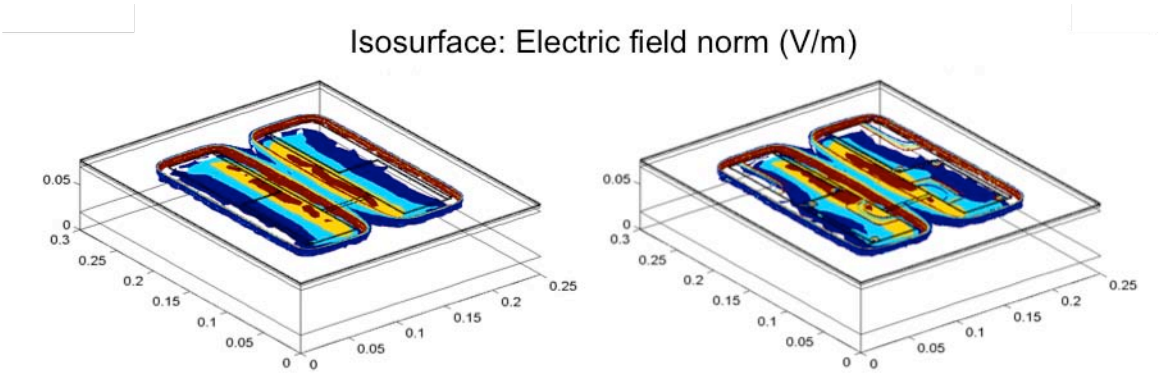


Figure 15: Conductive electrodes with wire. 3D ISO-surface plots of the electric field in the tissue with varying electrode conductivity. ISO-surfaces are plotted for electric field values of 10 V/m, 12 V/m, 14 V/m and 16 V/m. Electrode conductivity of 1 MS/m (left), 10μS/m (right).

2.2.3 Validating how leakage of stimulation gel influences the stimulation effect in the muscle tissue - or validating the use of conductive gels.

A film of conductive gel is traditionally applied under the electrodes in order to secure good electrical contact between the electrodes and the skin. Recently, however, more and more clinics use a thin layer of tap water - not related to increase conductivity, but rather to establish an air-free contact between the electrode and the skin. Air spaces may actually cause small sparlings that feel very unpleasant for the patient. Anyway, after about 20 minutes of pre-heating without stimulation, the patients will start to sweat and excrete salt into the thin water film and thus increase the conductivity, without the same risk for leakage as when the electrodes are greased with gels. In order to validate the leakage of gel from the electrode area to the gap between the electrodes we performed simulations.

In the simulation we varied the amount of gel spillage by varying the length of the area for which there is leakage (from 2 mm to 30 mm), and the gap between the two parts of the gel leakage (from 4% of the length between the electrodes covered with gel to all of the length between the electrodes covered with gel). We further assumed that the gel leakage is symmetric around the centerline between the two electrodes. An electrode conductivity of 40 S/m was applied in all simulations, and the conductivity of the gel was set to be 6.2 S/m.

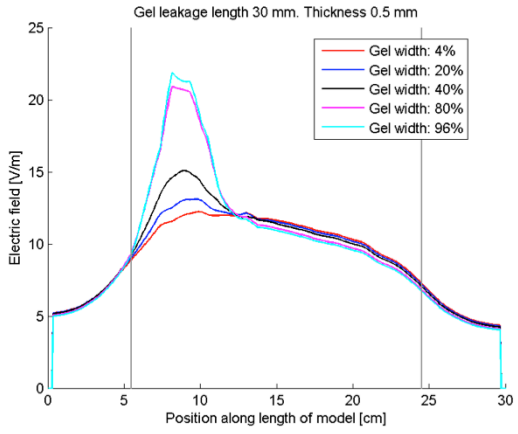


Figure 16: Variation of electric field in the muscle tissue with amount of gel leakage.

3 Conclusion

The current study reveals that COMSOL Multiphysicsnot only provides an excellent tool for choosing design but also for validating designs because of the fact that the software allows us to study functional implications of our design - a feature oneotherwise only can access in usability and clinical studies, which cannot take place before finishing safety tests (ISO 60601) of the prototype.

Furthermore COMSOL Multiphysicshas also been an excellent tool to minimize the use of laboratory animals complying with the European Directive 86/609/EEC.

4 References

1. Alexander, K., Clarkson, J., Bishop, D., and Fox, S. (2001). *Good Design Practice for Medical Devices and Equipment - A Framework*. The University of Cambridge Institute for Manufacturing, Mill Lane Cambridge CB2 1RX.ISBN 1-902546-08-3. pp 67.
2. Bronzino, J. D. (2000). The biomedical engineering handbook, Springer.
3. Carrara N, (2010). Dielectric Properties of Body Tissues: Webpage: http://niremf.ifac.cnr.it/tissprop/.
4. COMSOL Multiphysics® (2014) http://www.comsol.com.
5. "Directive 86/609/EEC on the protection of Animals used for Experimental and other scientific purposes" European Commission, 1986. Retrieved February 8, 2007
6. Drizdal, T., Togni, P., Visek,L. and Vrba, J. (2010). Comparison of Constant and Temperature Dependent Blood Perfusion in Temperature Prediction for Superficial Hyperthermia. *Radioengineering* 19, no. 2: 281.
7. Duck, F.A., (1990). *Physical properties of tissue: a comprehensive reference book*, Academic Pr.
8. EN 14971 (2001)
9. EN 46001 (1996)
10. EN 724(1994)
11. European Council Directive 93/42/EEC (1993)
12. Fiala, D., Lomas, K.J. &Stohrer, M., (1999). A computer model of human thermoregulation for a wide range of environmental conditions: the passive system. *J ApplPhysiol,* 87(5), 1957-1972.
13. GMP Regulations for Medical Devices (1976)
14. ISO 13485 (2003)
15. ISO 14971 (2007)
16. ISO 60601 (2008)
17. ISO 9001 (1944)
18. Kocbach J., Folgerø K., Mohn L., Brix O. (2011). A Simulation Approach to Optimizing Performance of Equipment for Thermostimulation of Muscle Tissue using COMSOL Vol 4, No 2, pp 9-33*Multiphysics., Biophysics &Bioeng. Letters*
19. Kocbach, J., Folgerø, K., Mohn, L., Brix, O. (2012). *Optimizing Performance of Equipment for Thermostimulation of Muscle Tissue using COMSOL Multiphysics*.Poster/AbstractComsol Conference Europe, Milan 2012.
20. Kreith, F. and Goswami, D. Y., (2005). *The CRC handbook of mechanical engineering*, CRC Press.
21. Kuhn, A., and Keller, T. (1999) 3D transient model for transcutaneous functional electrical stimulation. In *International functionalelectricalstimulationsocietyconference*, page 385–387.
22. Matweb, (2010), Webpage: Summary of similar materials in the MatWeb database for the category "Silicone Rubber", accessed October 28th 2010. http://www.matweb.com/search/DataSheet.aspx?MatGUID=cbe7a469897a47eda563816c86a73520
23. Medical Device Amendments (1992)
24. Nelson, R. M., Currier, D. P. and Hayes, K. W. (1999). *Clinical Electrotherapy*. 3rd. Edition, Appleton & Lange
25. Pennes, H. H., (1948). Analysis of tissue and arterial blood temperatures in the resting human forearm. *Journal of Applied Physiology*, 1(2), 93.
26. Primasil Silicones Ltd Data Sheet, (2010). Heat Cured Conductive Silicone Rubber (HCR), Type PR 610 / 60.
27. Quality System (QS) Regulation/Medical Device Good Manufacturing Practices (1997)
28. Safe Medical Devices Act (1990):
29. US FDA(1996)
30. Ward, A. R. (1984), Electrode Coupling Media for Transcutaneous Electrical Nerve Stimulation, *Australien J Physiotherapy*, 30, 3:82-85
31. Watlow (2010). Watlow Application Guide, Reference Data, Physical Properties of Solids, Liquids and Gases, Available from http://www.watlow.com/reference/files/nonmetallic.pdf, Accessed September 2010

Marketing and Neuroscience: as electroencephalographic tools could help to design and to analyze commercial advertising campaigns

Giovanni Vecchiato[*,**], Anton Giulio Maglione[**], Patrizia Cherubino[**,****], Ilenia Graziani[*], Arianna Trettel[**], Wanzeng Kong[*****], Fabio Babiloni[*,**]

[*] Dept. Physiology and Pharmacology, "Sapienza" University, Piazzale Aldo Moro, 5, 00185, Rome, Italy
[**] BrainSigns srl, Via Sesto Celere, 7/C, 00152 Rome, Italy
[***] Dept. Anatomy, Histology, Forensic Medicine and Orthopedics, Piazzale Aldo Moro, 5, 00185 "Sapienza" University, Rome, Italy
[****] Dept. Economics and Marketing , "IULM" University, Via Carlo Bo, 1, 20143 Milan, Italy,
[*****] College of Computer Science, Hangzhou Dianzi University, Hangzhou, China

giovanni.vecchiato@uniroma1.it, antongiulio.maglione@uniroma1.it, patrizia.cherubino@brainsigns.com, Ilenia.graziani@brainsigns.com, arianna.trettel@brainsigns.com, kongwanzeng@hdu.edu.cn, fabio.babiloni@uniroma1.it

Keywords: Neuromarketing – TV commercial – Cognition – Emotion

Abstract: Neuromarketing is the multidisciplinary field of research whose aim is to investigate the consumer's to commercial advertisements or to the product appreciation through the use of a neuroscientific perspective. In fact, there are high hopes that neuroimaging technology could solve some of the problems that marketers face. In particular, the neuroscience field is thought to be able to reveal information about consumer preferences which are unobtainable through conventional methods, including the administration of questionnaires to the consumers. Thanks to the high resolution electroencephalography (hr-EEG) tools, we were able to track the electric brain activity and correlate the spatiotemporal variations of such as signals with events of declared memorization and pleasantness. The present paper aims at illustrating some applications of electrical neuroimaging for the evaluation of marketing stimuli. The proper use of these methodologies can provide information related to cognitive and emotional aspects of persons watching TV advertisements. Evidences of this research show how EEG methodologies could be then employed both to better design new products that marketers are going to promote, as well as to analyze the global impact of video commercials broadcasted on TV.

1 INTRODUCTION

In the last two decades, different techniques and devices have been developed to measure thoughts and feelings by measuring the brain activity through the collection of the hemodynamic or electromagnetic signature of activated neural networks. Since last 10 years, unsolved problems and questions related to the evaluation of economic transactions, reached the neuroscience labs. Therefore, neuroscience researchers began to cooperate with economists in order to evaluate the brain activity during the generation of economic value judgments. What can cognitive neuroscience teach economics on the base of these last decades of common research? Although humans are definitely capable of conscious deliberation, many, if not most, economically relevant decision processes are characterized by certain other features such as automatic, fast and effective cognitive processes, which are not under direct volitional control (Bargh and Chartrand, 1999). Second, they are under the influence of unrecognized and finely tuned affective mechanisms, which often play a decisive role in action (Damasio et al., 1996; Davidson et al., 1990; Panksepp, 2004). Third, many of these processes have been shaped by evolution in order to serve social purposes (Adolph, 2003; Brothers, 1990; Cacioppo, 2002). Thus, decision-making and evaluation in economic contexts will be influenced by mechanisms dedicated to social interaction (Braeutigam et al., 2004). Hence, the incorporation of neuroimaging into the decision-making sciences has generated the interest not only of economists but also of marketing scientists. The line of reasoning was that if the neuroscientist could "image" the brain at work during the perception of particular concepts of value for the economists, such as the "perceived value" of choices to be made, then the same could be attempted for concepts useful in marketing. In fact, images of activated brain areas during the appreciation

of concepts, such as the "brand" or simply during the observation of advertisements, was attempted and successively largely explored by different brain imaging devices. Todays, there are high hopes that neuroimaging technology could solve some of the problems that marketers faced. In particular, a prominent hope is that neuroimaging will help them to streamline marketing processes and save money if commercial messages would be better broadcasted. Another hope is that neuroimaging will reveal information about consumer preferences that is unobtainable through conventional methods. In this context, the most popular brain imaging method adopted in the neuromarketing field is the functional Magnetic Resonance Image (fMRI), a technique that returns a sequence of images of the cerebral activity by means of the measure of the cerebral blood flow. It is very well known that the hemodynamic measurements of the brain activity allow a level of localization of the activated brain structures on the order of few cubic millimeters, being capable to detect activations also in deep brain structures. However, the lack of time resolution, due to the intrinsic nature of the measured signal, makes the fMRI unsuitable to follow the brain dynamics on the base of its sub seconds activity. However, there are other brain imaging techniques that allow to follow on a millisecond base the brain activity during the exposition to relevant marketing stimuli, such as electroencephalography (EEG), although characterized by a spatial resolution of square centimeters. To overcome this problem, high-resolution EEG technology (hrEEG) has been developed to enhance the poor spatial information content of the EEG activity in order to detect the brain activity with a spatial resolution of a square centimeter and the unsurpassed time resolution of milliseconds (Babiloni et al., 2004). In addition, EEG devices are also relatively inexpensive, robust and even wearable by the subject, making such technology of interest for the evaluation of marketing stimuli. Nowadays, researchers are attempting to investigate the signs of the brain activity correlated with an increase of attention, memory and emotional engagement during the observation of commercial advertisements (Vecchiato et al., 2010; Langleben et al., 2009). In fact, indirect variables of emotional processing could be gathered by tracking variations of the activity of specific anatomical structures linked to the emotional processing activity in humans, such as the pre- and frontal cortex (PFC and FC respectively; Davidson, 2004). The PFC region is structurally and functionally heterogeneous but its role in the generation of the emotions is well recognized. Specifically, findings suggest that the left PFC is an important brain area in a widespread circuit that mediates appetitive approach, while the right PFC appears to form a major component of a neural circuit that instantiates defensive withdrawal (Davidson, 2004; Balconi et al., 2009). In addition, it is very well know the role of the frontal areas in cognitive processes such as memory and attention in complex tasks (Werkle-Bergner et al., 2006; Klimesch, 1999). Moreover, by monitoring autonomic activity such as the heart rate (HR) it is possible to assess the emotional state of the subject (Montano et al., 2009).

The way in which such amount of information from neuroscience could be useful to the marketing studies is suggested by the well-known fact that at least the 70% of the new products launched worldwide (including cars, shoes, clothes etc.) failed within the first six months. This happens simply because people are not saying (or are not able to say) the truth when they are interviewed with respect to the product they have tasted or watched previously. There are a lot of reasons for that behavior, but this disability to proper report the own feelings related to the product is usually a huge obstacle to the correct collection of information from consumers, at the base of solid marketing actions. Thus, the application of neuroscience-based methodologies allows to the researchers to gain information not available otherwise related to the unconscious and spontaneous reaction of the consumers in front of the product stimuli. Of course such kind of information is just a part of the whole story related to the acquisition of consumer's habit in front of the product.

In the present context, the object of our research was to measure and analyze the brain activity and the emotional engagement occurring during the "naturalistic" observation of TV commercials. Hence, the final goal was to link significant variations of the EEG and autonomic variables with cognitive and emotional reactions to the presented ads. In order to do that, different indexes were employed to summarize the performed measurements, later used in the statistical analysis. Particularly, the procedure of the experimental task consists in observing a documentary in which a series of commercial video-clips are inserted. The experimental subjects were told to pay attention to the movie they would have watched, not aware that an interview would be held within a couple of hours from the end of the recording. In the interview, subjects were asked to recall commercial clips they remembered and to score them according to the degree of pleasantness they perceived. Hence, the dataset has been divided into several subgroups in order to highlight differences between the cerebral activity related to the observation of the remembered and forgotten ads, and those between the liked

against the disliked commercials. Finally, the experimental questions of the present study were the following: are there particular EEG activities correlating with the memorization and the perceived interest related to the observed TV commercial? Are there particular cerebral and autonomic indexes describing the emotional state experienced while watching the TV commercial? The interest for the marketing applications could derive from the following situation: a multinational company has to decide the way to promote in an Asiatic country a particular mobile phone device. There are two particular realizations of the advertising spot, one with the western actors and the other with the eastern actors. All the two spots have identical plot and music. Explicit interview with a sample of customers reported no particular preferences for one or other advertising. The idea is to use the measurement of the cerebral and body marker variables to have more information on which one of the particular spot has to be chosen by the company.

2 METHODS AND RESULTS

Twenty healthy undergraduate students of the Hangzhou Dianzi University (Hangzhou, Zeijiang, China) have been recruited for this experiment (age 22.95 ± 1.09; ten males). The task consisted in watching a ten minutes long documentary in which we inserted one advertising break, in the middle of the movie, formed by six TV commercials as experimental stimuli. The advertisements selected are related to three different international brands advertising sport equipment, hi-tech products and clothes, respectively. For each brand we have chosen a western and an eastern version of the clip containing no speech but advertising the item just by images and sounds. The reason why we made this decision lies in obtaining marketing messages which are understandable by both eastern and western populations. Randomization of the occurrence of the commercial videos within the documentary was made to remove the factor “sequence” as possible confounding effect in the following analysis.

In the following we are going to compare the neurophys-iological activity of experimental subjects while watching two versions of a hi-tech advertisement. The plot of these two commercials is the same: a father is far from home and celebrates his daughter by means of the advertised smartphone. Hence, it has been possible to define, for both commercials, seven segments of interest defined as follow-ing: Testimonial (comprising the whole ad except the expo-sition to the brand), Brand (exposition to the brand), Dad (scenes in which the only father is presented), Mum + Daughter (both mother and daughter are on the screen), Dad on smartphone (scenes in which the father is seen through the smartphone), Mum + Daughter on smartphone (scenes in which both mother and daughter are seen through the smartphone), Daughter (the only daughter appears on the TV screen). For each segment we compared the mean cerebral and heart rate activity in both advertisements as illustrated in the following.

Informed consent was obtained from each subject after explanation of the study, which was approved by the local institutional ethics committee.

All subjects were comfortably seated on a chair in front of a computer screen showing the experimental stimulus. We collected the EEG activity at a sampling rate of 256 Hz while the impedances kept below 5 kΩ by means of the gUSBamp amplifier (g.Tec medical engineering GmbH). The montage we employed refers to the 10-10 International System with the following recording channels: Fpz, AF3, AF4, F3, Fz, F4, T7, C3, Cz, C4, T8, P3, Pz, P4, Oz.

Raw EEG traces were first band pass filtered (hp=2 Hz; lp=30 Hz) and the Independent Component Analysis (ICA) was then applied to detect and remove components due to eye movements, blinks, and muscular artefacts. The extra-cerebrally referred EEG signals have been transformed by means of the Common Average Reference (CAR) and the Individual Alpha Frequency (IAF) has been calculated for each subject in order to define the frequency bands of interest according to the method suggested in the scientific literature (Klimesch, 1999). In particular we defined the following three frequency band: theta (IAF-6, IAF-4) i.e. theta ranges between IAF-6 and IAF-4 Hz, lower alpha (IAF-4, IAF) and upper alpha (IAF-2, IAF+2). EEG traces were then segmented to extract and analyse the cerebral activity during the observation of TV commercials. Each

EEG trace has been band pass filtered in order to isolate the spectral components in the theta, lower alpha and upper alpha bands from the whole EEG spectrum.

The filtered traces have been employed to calculate the Global Field Power (GFP; Lehmann D. and W. Skrandies, 1980). Since for the phenomena we would like to investigate a clear role of the frontal areas have been depicted (Davidson, 2004; Klimesch, 1999; Werkle-Bergner et al., 2006) we used the frontal electrodes to compute the GFP indexes used in the following of this study. In order to summarize the properties of the cerebral activation for the analysed ads we used the theta, lower alpha and upper alpha bands to define the Memorization, Attention and Pleasantness indexes, respectively (Vecchiato et al., 2010; Davidson, 2004; Klimesch, 1999; Werkle-Bergner, 2006). The filtered EEG traces were subjected to the computation of the GFP by taking into account the signals coming from the following frontal and prefrontal electrodes of the 10-10 International System: AF3, F3 to compute the Memorization Index (MI); Fpz, AF3, F3, AF4, F4, Fz to calculate the Attention Index (AI); homologous channels AF3, AF4, F3, F4 to evaluate the Pleasantness Index (PI). As to the Attention Index, we reversed the GFP waveform in order to have the activity of de-synchronization pointing up. As far as concern the Pleasantness Index, it has been defined by taking into account the frontal EEG asymmetry's theory by Davidson (2004) as already investigated in a previous study (Vecchiato et al., 2011). Hence, the formula defining the PI is the following:

$$PI = GFP_{dx} - GFP_{sx} \quad (1)$$

Where the GFP_{dx} and GFP_{sx} stand for the GFP calculated among right (AF4, F4) and left (AF3, F3) electrodes, respectively. The GFP signals of each subject have been averaged

The hrEEG technologies allowed to analyze the temporal trend of the cortical activities thanks to a high temporal and spatial resolution, distinguishing changes of activation of cortical areas by means of a graphical representation on an average brain model. The reconstruction of the cortical activity led to highlight the cerebral regions that were significantly activated when compared to the observation of the documentary, frame by frame. Statistical cortical spectral maps returned that the theta band activity during the observation of the TV commercials that were remembered is higher and localized in the left frontal brain areas when compared to the activity elicited by forgotten advertisements. Similar increase of the alpha rhythms occurred during the observation of advertisements that were judged pleasant when compared with the others (Davidson, 2004). The EEG signals have been also used to calculate the Global Field Power (GFP), an average measure of the brain activity (Davidson, 2004), to describe both cognitive and emotional variables to define the Memorization (MI), Attention (AI) and Pleasantness Index (PI). The percentage of spontaneous recall is linearly correlated with the MI values (R^2=0.68, $p<0.01$). The highest values of spontaneous recall correspond to MI values over average. In fact, in this case the percentage reaches the value of 33% when the AI is under average and the value of 41% when both MI and AI are over average (Vecchiato et al., 2012). As to the PI, the de-synchronization of left alpha frontal activity is positively correlated with judgments of high pleasantness (Vecchiato et al., 2011). In addition, the heart rate activity elicited during the observation of the TV commercials that were remembered or judged pleasant is higher than the activity during the observation of commercials that will be forgotten or were judged unpleasant (Vecchiato et al., 2012). The usefulness of such as indexes can be appreciated in the following figure which shows and compares, for a specific experimental group, the variation of AI and PI for two TV commercials. Particularly, cerebral data in the picture refer to twenty healthy undergraduate students of the Hangzhou Dianzi University (Hangzhou, Zeijiang, China) who watched an eastern and a western version of the same story-board of a TV commercial. The advertisements selected are related to an international brand selling hi-tech products. The two clips contain no speech but advertise item just by images and sounds with asian or with european actors. The plot of these two commercials is the same. Hence, it has been possible to define, for both ads, the same seven segments of interest describing different parts of the commercial.

As to the AI, it is possible to observe that all segments are characterized by statistically similar average values of GFP. These results suggest that for the analyzed subjects the two versions of the advertisement elicit the same level of attention. As far as concern the PI, we can observe that the eastern TV commercial is perceived more pleasant with respect to the western one. The difference of the PI are statistically significant in each segment of interest except for the Brand one, where on the screen is presented the same logo for both

advertisements. Present findings suggest the eastern population investigated is more attracted by actors and situations they perceive more familiar with respect to ones presented in the western version of the analyzed TV commercial. This could be of help for marketers since it seems to be important to adapt the commercial campaign according to the country in which it has to be promoted.

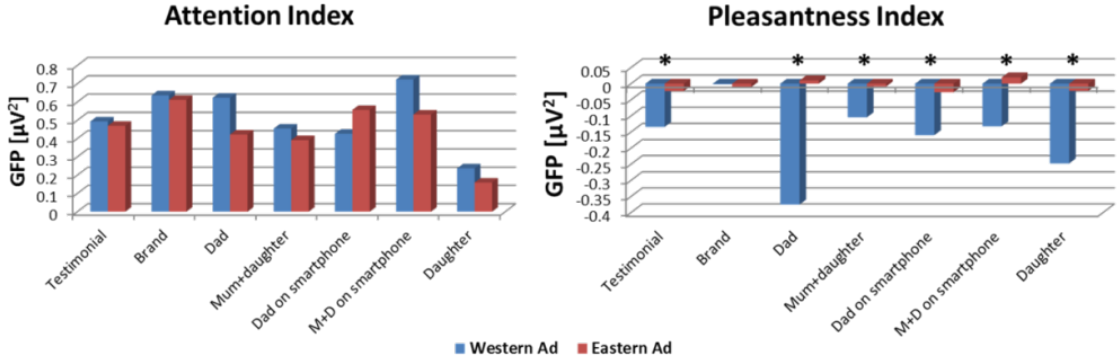


Figure 1. Figure presents the average values of GFP for the Attention Index (AI, left side) and Pleasantness Index (PI, right side), respectively. Each column refers to a single segment of interest showing the values of the cerebral variables for both western (blue) and eastern (red) advertisement. As to the AI No statistical difference, among the experimental conditions have been found. Instead, the differences of the Pleasantness Index are statistically significant for each segment except the Brand one (as the symbol * indicates; Student's t, $p<0.05$, Bonferroni corrected).

Overall, results underline that properties of the EEG rhythms, collected during the observation of TV advertisements, are linked with the overt preferences of the observers in terms of cognition and emotion. They can be used to generate metrics that automatically point to parts of the examined commercials that are congruent from the emotional and the cognitive point of view. These information could be used "a posteriori" to redraw the advertisement in order to highlight the pleasant frames while depressing the unpleasant ones, according to the population target. Finally, these tools allow to analyze the cognitive and emotional processes dynamics thanks to the high temporal resolution.

3. CONCLUSIONS

Marketers are excited about the use of brain imaging for marketing purposes mainly for two reasons. First, they hope that neuroimaging will provide a more efficient trade-off between costs and benefits. This hope is based on the assumptions that people cannot fully articulate their preferences when asked to express them explicitly, and that consumers' brains contain hidden information about their true preferences. Such hidden information could, in theory, be used to influence their buying behaviour, so that the cost of performing neuroimaging studies would be outweighed by the benefit of improved product design and increased sales. In theory, at least, brain imaging could illuminate not only what people like, but also what they will buy. Thus far, this approach to neuromarketing has focused on this post-design application, in particular on measuring the effectiveness of advertising campaigns. The second reason why marketers are excited about brain imaging is that they hope it will provide an accurate marketing research method that can be implemented even before a product exists. The assumption is that neuroimaging data would give a more accurate indication of the underlying preferences than data from standard market research studies and would remain insensitive to the types of biases that are often a hallmark of subjective approaches to valuations. If this is indeed the case, product concepts could be tested rapidly, and those that are not promising eliminated early in the process. This would allow more efficient allocation of resources to develop only promising products. From the marketing researchers point of view, there is the hope that these brain imaging techniques will provide an efficient trade-off between costs and benefits of the research. Improving the quality of the marketing messages allow the industries to lose less money in the production of inefficacy or incorrect advertisements and finally help them to match better the

demands of people related to the good to be advertised. The use of neuroimaging tools in the evaluation of the commercial ads help to reduce part of the money that were wasted in advertisement industry.

ACKNOWLEDGEMENTS

This work has been supported by the Italian and Chinese Minister of Foreign Affairs, in the framework of the bilateral Italy-China research program called “NEUROPREDICTOR”.

REFERENCES

Adolphs R. (2003). *Cognitive neuroscience of human social behaviour*. Nature Reviews. Neuroscience, 4(3), 165-178.

Babiloni F., Mattia D., Babiloni C., Astolfi L., Salinari S., Basilisco A., Rossini P.M., Marciani M.G., Cincotti F. (2004). *Multimodal integration of EEG, MEG and fMRI data for the solution of the neuroimage puzzle*. Magn Reson Imagin.vol. 22, no. 10, pp. 1471-1476.

Balconi M., Brambilla E., Falbo L. (2009). *Appetitive vs. defensive responses to emotional cues. Autonomic measures and brain oscillation modulation*. Brain Res. 2009 Nov 3; 1296:72-84. doi: 10.1016/j.brainres.2009.08.056. Epub 2009 Aug 21.

Bargh, J.A., Chartrand, T.L. (1999). *The Unbearable Automaticity of Being.* American Psychologist, 54(7), 462-479.

Braeutigam S., Rose, S.P., Swithenby, S.J., Ambler. T. (2004). *The distributed neuronal systems supporting choice-making in real-life situations: differences between men and women when choosing groceries detected using magnetoencephalography*. The European Journal of Neuroscience, 20(1), 293-302.

Brothers L. (1990). *The social brain: A project for integrating primate behavior and neurophysiology in a new domain*. Concepts in Neuroscience, 1, 27-51.

Cacioppo, J.T. (2002). Foundations in Social Neuroscience illustrated edition. The MIT Press.

Damasio, A.R., Damasio, H., Christen, Y. (1996). *Neurobiology of Decision-Making*. 1° ed., Springer.

Davidson RJ. (2004). *What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research*. Biol Psychol. 67(1-2):219-33.

Davidson RJ., Ekman P., Saron C.D., Senulis J.A., Friesen W.V. (1990). *Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I*. Journal of Personality and Social Psychology. Vol.58, No. 2, 330-341.

Klimesch W. (1999). *EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis*. Brain Res Rev, vol. 29, no. 2, pp. 169-95.

Langleben DD, Loughead JW, Ruparel K, Hakun JG, Busch-Winokur S, Holloway MB, Strasser AA, Cappella JN, Lerman C. (2009). *Reduced prefrontal and temporal processing and recall of high "sensation value" ads*. Neuroimage. 15;46(1):219-25.

Lehmann D. and W. Skrandies, *Reference-free identification of components of checkerboard-evoked multichannel potential fields*, Electroencephalogr Clin Neurophysiol 48 (1980), pp. 609–621.

Montano N., Porta A., Cogliati C., Costantino G., Tobaldini E., Casali K.R., Iellamo F. (2009). *Heart rate variability explored in the frequency domain: a tool to investigate the link between heart and behavior*. Neuroscience and Biobehavioral Reviews, 33(2), 71-80.

Panksepp, J. (2004). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. 1° ed., Oxford University Press, USA.

Vecchiato G., Kong W., Maglione A.G., Wei D. (2012). “*Understanding the Impact of TV Commercials: Electrical Neuroimaging*”. IEEE Pulse, Vol.3 (3), Page(s): 42-47.

Vecchiato G., Toppi J., Astolfi L., De Vico Fallani F., Cincotti F., Mattia D., Bez F., Babiloni F. (2011). *Spectral EEG Frontal Asymmetries Correlate with the Experienced Pleasantness of TV Commercial Advertisements*. Medical & Biological Engineering & Computing.

Vecchiato G., Astolfi L., De Vico Fallani F., Cincotti F., Mattia D., Salinari S., Soranzo R., Babiloni F. (2010). *Changes in brain activity during the observation of TV commercials by using EEG, GSR and HR measurements.* Brain Topography 23, n°. 2: 165-179.

Werkle-Bergner M., Muller V., Li S.C., Lindenberger U. (2006). *Cortical EEG correlates of successful memory encoding: implications for lifespan comparisons*. Neuroscience and Biobehavioral Reviews, 30(6), 839-854.

Biophysical Forces involved in Carcinogenesis

Maria Grazia Masiello[1,2], Alessandra Cucina[2], Mariano Bizzarri[3*]

1Department of Clinical and Molecular Medicine, "Sapienza" University of Rome, p.za Sassari 3, 00161 Rome, Italy
2Department of Surgery "PietroValdoni", "Sapienza" University of Rome, via A. Scarpa 14, 00161 Rome, Italy
3Department of Experimental Medicine, "Sapienza" University of Rome, Viale Regina Elena 324, 00161 Roma, Italy

Keywords: *Cancer Biophysics, Cytoskeleton, Nucleoskeleton, Mechanotransduction*

Abstract

Cancer initiation, progression and spreading is widely influenced by a wide variety of physical factors. To help tackle this challenge, several research groups have undertaken to investigate the role physical cues play in cancer biology, leading hence to a rediscovery of "cancer biophysics". Physical forces are thought to modulate cell and tissue behavior mainly by means of mechanotransduction (trough which physical forces are transduced to biochemical and genetic pathways) involving surface tension, stiffness, extracellular matrix (ECM) remodelling, cytoskeleton (CSK) and nucleoskeleton (NSK) rearrangement. Mechanotransduction contributes by that way to influence the morphology cells and tissue acquire, and thereby a wide array of shape-dependent processes (including proliferation and apoptosis).

* *Corresponding Author: Prof. Mariano Bizzarri
e-mail: mariano.bizzarri@uniroma1.it ; website: www.sbglab.org

Introduction

Large-scale cancer genomics, proteomics and RNA-sequencing efforts are currently mapping in fine detail the genetic and biochemical alterations that occur in cancer. However, it is becoming clear that gene-proteins interaction does not represent the whole picture. Needless to say that genetic factors, though necessary, do not work in isolation and they are significantly shaped and influenced by both physical and epigenetic cues. That complex interacting network ("interactome") is hardly to get known, given the difficulty in integrating/interpreting those data and in translating them into treatment. On the other hand, this complexity is likely to be at the root of the failure of current target-based treatments, largely grounded on a rigid reductionist, genetic-centered model.

Cancer initiation, progression and spreading is widely influenced by a wide variety of physical factors. To help tackle this challenge, several research groups have undertaken to investigate the role physical cues play in cancer biology, leading hence to a rediscovery of "cancer biophysics"[1].

Physical forces are thought to modulate cell and tissue behavior mainly by means of mechanotransduction (through which physical forces are transduced to biochemical and genetic pathways) involving surface tension, stiffness, extracellular matrix (ECM) remodelling, cytoskeleton (CSK) and nucleoskeleton (NSK) rearrangement. Mechanotransduction contributes by that way to influence the morphology cells and tissue acquire, and thereby a wide array of shape-dependent processes (including proliferation and apoptosis). Tissue organization as well is significantly affected by gradient's force and different field's effects (as those provided by gravity and electromagnetic field). Yet, physical cues should not be considered as "efficient causes", within the Aristotelian meaning, but as 'constraints', i.e. factor able in driving the system toward discrete, different issues.

Where do the restraints come from in biological systems? Clearly, the immediate environment of the system is one source of restraint. Cells are restrained by the tissues and organs they find themselves in, meanwhile internal constraints belong principally to the architecture taken by CSK and NSK. Even these restraints though would not exhaust the list. Indeed, organisms are also a product of their evolutionary history, that is to say the integrated sum of their interactions with past environments. Overall, this kind of complexity account for the fact that cancer understanding cannot be restricted to the analysis of protein and gene circuits.

The Mechanotransduction apparatus

The ability of a cell to sense mechanical properties and changes that result in varying cellular response can be divided into three major steps – mechanosensing, mechanotransduction and mechanoresponse. Mechanosensing is a term used to define the ability of a cell to sense the mechanical properties of the environment. A compelling body of evidence has evidenced that cells are shaping their architectural skeleton – the cytoskeleton (CSK) – in order to acquire a "tensegrity" architecture. Performance of that structural system depends on tensile prestress for its mechanical stability and enables cells to sense and respond to mechanical inputs. That model has been proposed first by R.B. Fuller to describe structures in which stability ("architectural homeostasis") relies on the "tensional integrity", that is a 'continuous, diffuse tension that allows a discontinuous compression' [2]. By coupling tensile and compressive forces, that architecture allows each structural component to be in a pre-tensed condition. By this way, tensegrity stabilizes cells shape and make them to be highly responsive to external perturbation driven by physical stimuli. Response to mechanical stress involves primarily a modification of the tensegrity structure, i.e. a change in the CSK conformation, and thereafter a set of different changes in protein conformation, kinase activation or protein clustering which lead to biochemical reactions as well as gene regulation [3](mechanotransduction). Ultimately, such effects end up in modifying cell shape, motility, proliferation and other complexes cell function (mechanoresponse).

Usually cells experience a different set of forces involving shear stress (where the force is applied parallel to the surface of the cell), compressive stress (when the stress is applied perpendicular to

the cell resulting in compaction) and tensile stress (when the tension exerted on the cell leads to an expansion)[4].The balance of internal and external stresses, sensed by a mechano-sensor (like focal adhesion, FA), drives the tensional homeostasis experienced by the cell and the subsequent up or down regulation of fundamental cellular functions such as proliferation, adhesion, migration or differentiation. Behavior and dynamics of focal adhesion is mainly modulate by the tyrosine phosphorylation and dephosphorylation of the various focal adhesion proteins. The most significant kinases involved in FA are focal adhesion kinase (FAK) and c-Src [5]. Upon inducing a mechanical stress on a cell, the focal adhesion proteins undergo enhanced tyrosine phosphorylation, and the extent a single cell spreads on a substrate can be correlated proportionally with the phosphorylation increase [6]. That process is highly dependent of the adhesion substrate and on its mechanical properties, from which ultimately the emerging stiffness depends [7,8].
Thus, the stress balance depends on both active elements (intracellular contraction or traction exerted by neighbouring cells) and passive components (cytoskeleton and ECM stiffness). Then, softening or hardening, i.e. remodelling of these visco-elastic components, directly affects the local tensional homeostasis.

Cell-matrix and cell-cell interactions

Cells are physically connected each other by specialized structures, and to the ECM anchoring scaffold.
ECM-anchorage is provided by integrins, a specific class of cell surface receptors, spanning trough the cell membrane and endowed with an extracellular and a cytoplasmic tail. Integrins are heterodimeric glycoprotein composed of α and β subunits. To date more than twenty integrins subtypes have been recorded, and that diversity enables cells to mediate anchorage to a wide array of ECM proteins (laminin, vitronectin, fibronectin, collagen fibers, fibrinogen). The cytoplasmic tail of integrins binds to different actin-binding proteins (talin, vinculin, paxillin,) that physically link actin to the CSK [9]. At the focal adhesion (FA) site multiple integrins and their intracellular binding partners cluster together to form anchoring complexes to which the terminal ends of contractile actomyosin filaments of the cytoskeleton are inserted. By this way FA experiences a pre-existing tension and may sense mechanical stresses, either applied externally to integrins or transmitted from the inner contractile CSK, by modifying its molecular assembly. Calcium influx is mediated by physical stress and in turn it plays a critical role in activating cell surface integrins, and in promoting their attachment to ECM proteins[10,11]. Within the FA framework, Cadherins play a very relevant role. Cadherins are calcium-sensitive proteins participating in the formation of lateral cell-to-cell junctions, a critical structure involved in maintaining tissue integrity [12]. Even more important, the cytoplasmic tail of Cadherin is connected to actin filaments through α-catenin and β-catenin, as well as by a few other specialized proteins, like vinculin [13]. By this way Cadherin-Catenin complexes may contribute in sensing mechanical cues at cell-to-cell junctions and therefore participate in tensional homeostasis [14].

Cytoskeleton and nucleoskeleton

The CSK, a complex framework of interconnected microfilaments and microtubules, provide the inner structure embedded into the cytosol of eukariotic cells [15]. Cytoskeletal filaments both generate and resist mechanical loads. In addition, cytoskeleton components contribute in shaping cell form and in assuring resilience to shape distortion. The CSK is a network of three major structural elements: microtubules, intermediate filaments, and microfilaments, each consisting of polymers of protein subunits [16]. Cells generate mechanical tension in their actin CSK and exert tractional forces on their adhesion to ECM. Changes in the balance of forces between cells and ECM, induce modifications in surface tension and in matrix properties (flexibility, stiffness, adhesivity). By this way physical cues can change cell shape and switch cells towards different phenotypic fates [17,18]. Protein complexes at the ECM-membrane-actin cytoskeleton junctions

(e.g., focal adhesions and fibrillar adhesions) provide physical connections between the extracellular and intracellular compartments, foster cell-to-cell connectivity, and eventually orchestrate actin cytoskeleton organization, cell shape modulation, and other fundamental cellular processes [19]. Early events in force detection mechanically induce cytoskeletal changes that result in biochemical signals to mechano-responsive pathways than ultimately regulate cell form [20].
Components of the CSK play a key role in motility, transport and cell division, providing essential scaffolding on which metabolic processes occur. Therefore cytoskeletal morphology is thought to be a valuable indicator of cell injury and functionality [21]. Inner cytoskeleton structure provides also 'privileged' pathways along which enzymes and substrates are coherently organized and oriented, in order to optimise their interactions [14]. Those biochemical activities are in turn deeply affected by shape changes and mechanical stresses that interact with the cytoskeleton architecture [22]. A key feature of the cytoskeleton is that it is in a state of isometric tension, which ensures that various molecular-scale mechanochemical transduction mechanisms proceed simultaneously and produce a concerted response. It is likely that at least one pivotal mechanism through which these complex behaviours are modulated is mediated by shape control on focal adhesion structures. Indeed, focal adhesion formation and organization are governed by both internal cytoskeletal and mechanics that result from large-scale changes in cell shape [23].
Mechano-transduction of physical forces along the cytoskeleton and the adhesion structures involves the nucleus organization as well. Because extracellular forces are transmitted to the nucleus, where they cause substantial deformations, it should be no surprise if these forces could directly or indirectly contribute to changes in chromatin structure, transcriptional activity and nuclear organization, given that the 3D- architecture of chromatin is a critical component of nuclear gene regulation [24]. "These features of living architecture are the same principles that govern tensegrity (tensional integrity) architecture, and mathematical models based on tensegrity are beginning to provide new and useful descriptions of living materials" [25].
Indeed, DNA in vivo is often sharply distorted away from the canonical Watson–Crick structure; different DNA sequences vary greatly in the ease with which such sharp distortions can be accommodated. Most of the eukaryotic genomic DNA is bent around histones to form nucleosomes. The capacity of the DNA sequence to undergo such distortion can influence the specific preferred locations for many of the nucleosomes, thus enabling specific gene sequences to be or not to be expressed. By that way changes in nucleoskeleton may modulate genomic activity [26].

Biophysics of the tumor microenvironment

By considering the origination and transformation of developmental mechanisms as an evolutionary problem in its own right we observe that non-genetic mechanisms are a major sources of morphological novelty in evolution. These mechanisms include a) interactions of cell metabolism with the physicochemical environment within and external to the organism, b) interactions of tissue masses with the physical environment on the basis of physical laws inherent to condensed materials, including electromagnetic forces, c) interactions among tissues themselves, according to an evolving set of rules. The forms and characters assumed by metazoan organisms originated in large part by the action of such processes [27].
Namely, physical forces sensed by cells act as driving forces in morphogenesis. Surface tension and cytoskeletal contraction concur in directing embryo cells differentiation [28]. When cultured on a hard surface, cells spread out and cytoskeletal contraction generates high levels of tensile forces that pull on the surface. These changes promote the differentiation of stem cells toward the osteoblast lineage. Over-expression of either Rho or Rho-associated kinase (Rock), which both stimulate contraction of the actin cytoskeleton, also promotes osteoblastic differentiation. On the other hand, cells can be kept round by preventing them from spreading, inhibiting actin polymerization, increasing cell density, or encapsulating them in a soft gel [29].

Therefore, tensional homeostasis is likely to be essential for normal tissue growth and differentiation, and increasing tissue rigidity by stiffening the matrix (fibrosis) or by elevating Rho signaling could induce cytoskeletal contractility to enhance integrin-dependent growth and destabilize tissue architecture. As such, conditions that induce tissue fibrosis (matrix stiffening) or deregulation of cell-matrix interactions may facilitate malignant transformation by increasing cell contractility and/or disrupting the normal cytoskeleton architecture.

Indeed, normal tissue structure is disrupted during tumor initiation and progression. The microenvironment becomes both mechanically and biologically active, highlighted by continuous and progressive remodelling of the tumor mass and the stromal compartments. Within the tumor mass, the tumor onset is associated to loss of cell and tissue polarity, and with alterations in the composition and organization of extracellular matrix components. Parallel modification in behavior and structure of via multiple stromal components (cells, ECM molecules, vascular architecture) have been also recorded. Tumor-associated stromal cells, including fibroblasts, myofibroblasts, endothelial cells, mesenchymal stem cells, inflammatory cells and immune cells are often recruited, locally differentiated or activated during different tumor development stages. These cells actively participate in tumor initiation by shaping ECM remodelling and tumor angiogenesis, inducing the release of several paracrine molecular factors and thus enhancing tumor growth and metastasis [30, 31]. In addition, the non-cellular components of the tumor stroma, suchas collagens, fibronectin, tenascin and proteoglycans, undergo dramatic changes which significantly modify both the stiffness and the compliance of the surrounding milieu. In turn, changes in microenvironment viscosity and stiffness deeply influence cell tensegrity, thus leading to a consequent change in its shape and responsiveness to physical cues [32] (Fig. 1).

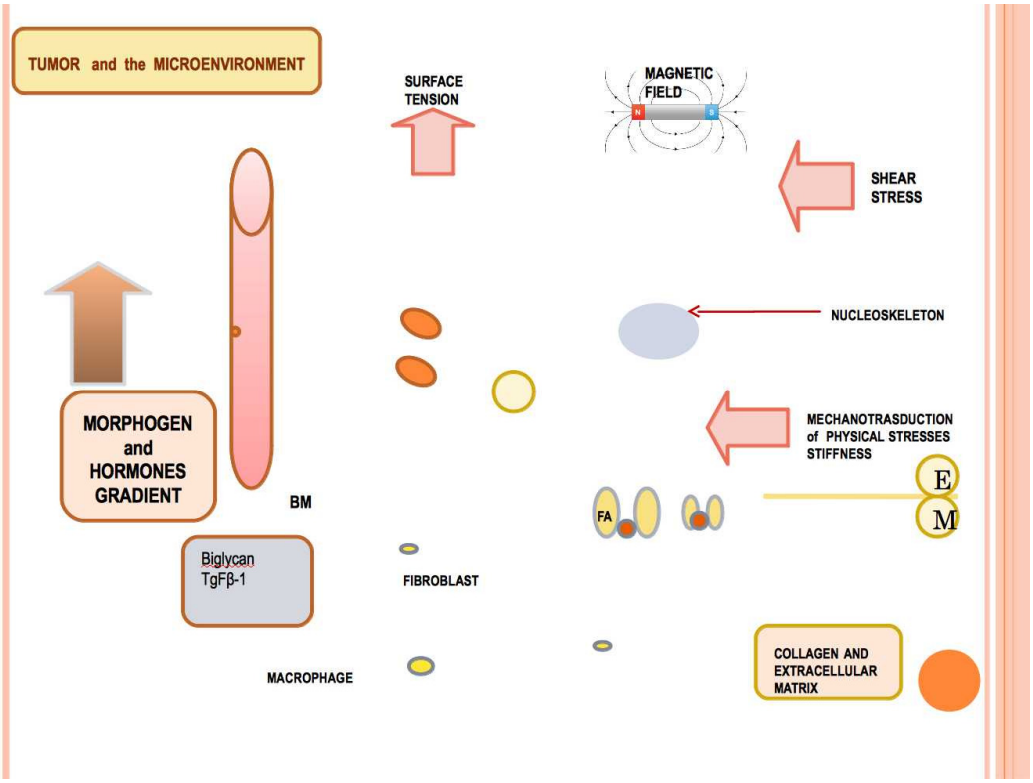


Fig. 1. Diagram representing biophysical forces and cues modelling the interactions among epithelial cells and their surrounding microenvironment. Legend: E, epithelium; M, Microenvironment; FA, Focal Adhesion site; BM, basal membrane.

The disruption of tissue structure and organization during tumor onset and evolution lead to unbalanced physical forces and altered material properties of each tumor component. However, as the cell-microenvironment interaction is highly dynamics, it is not surprising that the mechanical state in a tumor also evolves as the tumor develops. Microenvironment selective changes occur in different pathological conditions that are thought to precede the overt tumor transformation. Indeed, breast hyperplasia typically involves the loss of normal cell polarization and disruption of cell-cell contacts and cell-ECM interactions, accompanied by increased matrix deposition, cell proliferation, and thickening of the basement membrane (BM) architecture [33, 34]. On the contrary, carcinoma *in situ*is characterized by active cell growth within an intact BM and interstitial ECM. It is worth of noting that an intact BM 'constrain' the tumor growth within certain boundaries given that both BM and stromal components exert a resistance against the expanding tumor mass [35, 36]. Yet, the ensuing ECM remodelling and stromal reaction, lead to changes in the mechanical properties of the cell-microenvironment system, that ultimately further increase cell-generated forces and cell tension [37]. In advanced carcinoma in situ tissue stiffness further increase as a consequence of increase in ECM density and interstitial fluid pressure [38]. Invasive tumors display even further loss in cell-to-cell adhesion, increased intracellular contractility (and hence motility), promoting thus tumor invasion and spreading through the BM and ECM, finally ending to the peri-vascular space [39]. Indeed, shear stress and microenvironmental stiffness play a pivotal role in maintaining normal cell architecture and in limiting cancer cells motility. In absence of shear stress connective tissue morphology is highly disorganized and presents as a random array of polygon-shaped cells that is typical for primary high density cultures; on the contrary, chondrocytes exposed to fluid-induced shear stress for 48 hours exhibited elongated spindle-shaped cell bodies that appear in close contact and in alignment with neighbouring cells, leading to a well-organized tissue structure [40].Loss of tissue organization as well as gradient of tissue interstitial forces may influence significantly tumor cell migration and extravasation [41]. Eventually, as different mechanical properties are associated to each organ/tissue, and since cells can selectively grow within specific substrates according to their physical properties, organ-specific mechanical properties are of significant relevance in influencing the preferential migration, attachment, and proliferation of cancer cells [42, 43].

Studies on tumor microenvironment are dating back even from 1940, when microenvironment was shown to suppress skin carcinogenesis induced by chemical carcinogens [44]. Even experiments done to demonstrate that a single or few mutated genes are needed to induce carcinogenesis, were obliged in recognizing that microenvironmental factors are mandatory required to promote oncogenesis at the tissue level [45, 46]. As case in point, experimentally tumors obtained by inoculating cells with oncogenic virus demonstrated that the context play a pivotal role in driving the neoplastic transformation. The tumorigenicity of polyoma virus-transformed BALB/C 3T3 cells in syngeneic mice depend on the microenvironment in which these cells were grown rather than on the content of the polyoma middle T oncogene [47]. Moreover, given that no specific genetic traits have been associated so far to the metastatic process, despite outstanding efforts to find a correlation among genome profile and cancer malignancy [48], increased attention has recently been deserved to the microenvironment thought to confer a 'metastatic phenotype' [49], according to the "seed and soil" hypothesis, first proposed by Stephen Paget [50].

Indeed, those preliminary investigations highlighted that even potent carcinogenic cues could be overridden by embryonic microenvironment [51], a finding that recently has received compelling confirmation. Cancer cells cultured in embryonic environment [52-54] or co-cultured in 3D-matrices with normal human cells can be indeed committed to apoptosis and differentiation, and eventually reprogrammed into normal phenotypes [55, 56]. Such effects have been ascribed to some, even unrecognized "signalling molecules", morphogens or other soluble factors provided by the morphogenetic embryonic field. Yet, similar results have been obtained by culturing cancer cells into 3D-matrices with normal human cells. Indeed, a matrix containing both type I collagen

and reconstituted basement membrane, and the presence of normal breast fibroblasts constitute the minimal permissive microenvironment to induce near-complete tumor phenotype reversion [57].
For a while, interest on immunologic [58] and angiogenic [59] aspects of tumor microenvironment hasovershadowed the contribution of the microenvironment on cancer initiation [60]. Yet, it is now firmly established that the microenvironment actively contributes to carcinogenesis already since the first stages, given that it profoundly influences many cellular processes like growth, differentiation, and apoptosis [61, 62]. Even subtle differences in ECM composition and stiffness, selectively foster or inhibits proliferation, by modulating regulatory molecules of the cell cycle and the early response genes [63-65]. The microenvironment regulates the transcription of genes associated with differentiating pathways [66, 67]and participates in shaping cells phenotypes, by modulating cell-stroma interactions and cytoskeleton architecture [68, 69]. Moreover, cell shape and microenvironmental cues may trigger programmed cell death signals, hence driving cells towards apoptosis [70, 71]. Changes in the microenvironment structure or composition frequently lead to tissue fibrosis, augmented collagen crosslink, and tissue stiffening, which have been associated to an increased risk of developing cancer [72, 73]. It is not trivial to recall that aging is associated to increase in both tissue stiffness and cancer incidence [74]. In turn, tissue fibrosis and modification of physico-chemical properties of ECM may likely influence tumor onset and progression by regulating soluble factors involved in inflammation [75] and angiogenesis [76].

Physical premises for a microenvironment-based theory of carcinogenesis.
In the last decades it became clear that cell behavior is far from being "controlled" by linear (digital) 'commands', but rather by complex networks of molecular interactions and biophysical cues, spanning across different levels of structural and functional organization. Those interactions are not only context-dependent (and thus cannot be understood by keeping cells in an inappropriate milieu, like that provided by 2D-cultures), but also they obey to a non-linear dynamics which make impractical modelling processes with more than two variables. The switching between different stable states (representing differentiated or pathological phenotypes), requires that the activity/expression of several "signaling" molecules change in concert: indeed, mitogen stimulation or phenotype reversion is associated with the simultaneously co-expression of hundreds of different transcription factors and multiple downstream genes [77, 78]. To achieve a state-transition, no individual point mutation is actually needed, and even a cumulative effect linked to mutation will occur only if a critical state of the system as a whole is reached. It is worth of noting that the transition beyond that critical point "may be prevented or reversed by simultaneously manipulating a number of factors in the extracellular medium" [79]. Indeed, If multiple molecular elements must be tuned simultaneously to change cell phenotype, then it should be hypothesized that only a stimulus outfitted with pleiotropic property would perform that task, mainly based on stochastic fluctuations that enable transition from one attractor (phenotype) to another: that model may explain the genome-wide adaptability to environmental changes without requiring specific molecular signaling transducers [80]. This may explain why switching in between different cell fates can be triggered by changes in extracellular matrix structure, by inducing cell shape modification, by adding aspecific chemical substances, electrical ion flows, magnetic or gravitational fields [81-83]. Overall, those factors shared a meaningful property, given that they are able to modify the morphogenetic field and the biomechanical features of the systems [84].
Cells sense and respond to external physical forces and changes in matrix mechanics by modulating their endogenous cytoskeletal contractility: indeed, the mechano-sensitivity of cells lies on the delicate force balance between the endogenous cytoskeletal contractility and external mechanical forces transmitted across the cell-ECM adhesions [85]. The force balance is transmitted across the mechanical continuum of ECM-integrin-CSK regulates integrin-mediated adhesion signaling (such as FAK and Src signaling), where the adhesion sites in which integrins provide the mechanical linkage between the ECM and the actin CSK. Exposure of cells to mechanical strain, fluid shear

stress, or plating cells on substrates with varying elastic moduli, will activate integrins, which promote recruitment of scaffold and signaling proteins to strengthen adhesions and to transmit biochemical signals into the cell. These mechanotransduction pathways establish positive feedback loops in which integrin engagement activates acto-myosin CSK contractility, which in turn reinforces adhesions. Thus, the level of CSK contractility generated inside the cell is directly proportional to the adhesion strength and the matrix elastic modulus and dictates the cellular responses of cells [86]. Moreover, the way a cell sense and respond to a biochemical input (like those provided by growth factors) mainly depends on the *physical state* of both cells and their microenvironment. For instance, TGFβ-1 exerts a 'dual' role on cancer cells, and that paradoxical behavior is well recognized as a challenging enigma that, still now, classic molecular biology has not been able to elucidate [87-90]; on the contrary, by referring data to a higher level of observation - i.e., when the cell-microenvironment interaction or the tissue level are keep in consideration - conflicting results end up being such, and paradoxes may likely find a compelling explanation. Indeed, soluble factors like TGFβ-1 may trigger opposite outputs depending on the tissue stiffness: under mechanically unloaded conditions (floating matrices), TGFβ-1 stimulated contraction directly as an agonist and indirectly by pre-activating cells to express the myofibroblast phenotype, whereas, under mechanically loaded conditions (stressed matrices), TGFβ-1 had no direct agonist effect on contraction [91].

Physical and biochemical changes occurring within the microenvironment will be eventually transmitted from the cytoskeleton to the nucleoskeleton, thus enabling the selective unfolding of chromatin [92]. The DNA is enveloped in histone proteins to form strand, further wrapped and folded. Gene switching (on/off) can proceed properly only if the appropriate section of chromatin is unpacked and exposed to the enzyme machinery. This physical rearrangement of the chromatin is mainly dependent on the tensional forces perceived by the cell-microenvironment system and further transmitted across the focal adhesion along the cyto-nucleoskeleton to the cell biochemical/genetic machinery. Therefore, different cytoskeleton arrangements will end up in activating different gene sequences, leading hence in triggering different biochemical pathways [93]. The balance between tensional forces and the cytoskeleton architecture modulates thereupon several complex cell functions like apoptosis, differentiation, proliferation, ECM remodelling and so forth. That model can help in understanding the "dual" role displayed by a lot of "signaling molecules", selective sensitivity to drugs [94], as well as why cancer cell behavior may proceed "irrespective" of their "mutated" genes [95]. That is precisely what means "to put the gene in the context", given that the cell response to molecular "signals" tightly lie on the response of individual cells to mechanical tension and to the specific microenvironment in which cells are embedded. To date, an overwhelming body of data has revealed that mechanical tension generated through molecular interactions within the cytoskeleton is critical for modulating growth factors activity [96, 97] and to dramatically influence cell form and function [98]. In turn, interactions between epithelia cells and microenvironmental components (namely stromal cells), change ECM composition as well as its biochemical-biophysical features [99].

Some experimental results have provided compelling evidence demonstrating the key role of the microenvironment in cancer initiation. Despite the presence of growth factor, normal cells cannot grow when they are free of adhesion to ECM [100], or if they are compressed into specific geometric space (i.e., only along a thin epithelial monolayer) [101]. Similarly, stimulated breast cancer cells cease to grow when are detached to their substrate in a microgravity field [102]. Therefore, an increase in "signaling molecules" alone cannot explain cell growth induction. Therefore, it seems that the physical interaction with the microenvironment enables cells to respond to soluble factors or genetic inputs. Even in autosomal dominant tumor predisposition syndromes, like neurofibromatosis-1 (NF-1), NF-1 inactivation results in increased astrocyte growth, but the augmented proliferation rate is unable to induce glioma formation [103]. To observe tumor formation in vivo, brain microglia carrying NF-1 heterozygosity are needed. In that model,

microenvironmental components drive the epithelial transformation, mainly by providing disruption of ECM integrity (through the enhanced release of hyaluronidase), and subsequent activation of the MAPK-pathway. As expected, inhibition of hyaluronidase release or microglia activation dramatically reduces mouse optical glioma proliferation in vivo [104].
Overall, those results highlighted how biophysical factors participates in promoting and shaping the carcinogenic process that can be considered as a "development gone awry" [105]. As recently recognized, "the physical laws and principles that define the behaviour of matter are essential for developing an understanding of the initiation and progression of cancer", thus providing "opportunities for new insights into long-lasting problems in cancer research" [106]. That premise, well grounded on experimental basis, represents another discontinuity point in respect to the SMT for which "biological-information" carried out by genes constitute the only (or the main) causative factor in driving cellular fate and behavior.

Conclusion

Relevance of physical and structural cues in biological processes has been highlighted by studies performed on 3-dimension (3D). Indeed, studies with 3D model systems have repeatedly identified complex interacting roles of matrix stiffness and composition, integrins, growth factor receptors, and signaling in development and cancer [107].
These insights suggest that plasticity, regulation, and suppression of these processes can provide strategies and therapeutic targets for future cancer therapy and stem cell engineering. Additionally, besides providing model systems for testing ideas and potential therapeutic interventions, they may also permit high-throughput drug screening on human tissues in vitro. Yet, by focusing on such aspects, we are witnessing a *renaissance* of biophysical studies aimed in shedding a different light on carcinogenesis [108]. This is worth of noting, given that this approach implies a shift from molecules to cells and tissues, deemed as the privileged level of observation, i.e. the level in which the observed phenomenon is thought to arise. By that way, biophysics paves the way for an epistemological 'revolution' (according to Kuhn's definition) [109] in cancer studies.

References

[1] Michor F., Liphardt J., Ferrari M. and Widom J. What does physics have to do with cancer? *Nat Rev Cancer*,11(9):657-670, 2011.
[2] Swanson R.L. Biotensegrity: a unifying theory of biological architecture with applications to osteopathic practice, education, and research-a review and analysis. *J Am OsteopathAssoc*,113(1):34–52, 2013.
[3] Bershadsky A.D., Balaban N.Q. and Geiger B. Adhesion-dependent cellmechanosensitivity. *Annu Rev Cell Dev Biol*, 19:677-695, 2003.
[4] Butcher D.T., Alliston T. and Weaver V.M. A tense situation: forcing tumour progression. *Nat Rev Cancer*, 9(2):108–122, 2009.
[5] Sastry S.K. and Burridge K. Focal adhesions: a nexus for intracellular signaling and cytoskeletal dynamics. *Exp Cell Res*, 261(1):25-36, 2000.
[6] Lin V.C., Ng E.H., Aw S.E., Tan M.G., Ng E.H. and Bay B.H. Progesterone induces focal adhesion in breast cancer cells MDA-MB-231 transfected with progesterone receptor complementary DNA. *MolEndocrinol*,14(3):348-358, 2000.
[7] Schmidt C., Pommerenke H., Durr F., Nebe B. and Rychly J. Mechanical stressing of integrin receptors induces enhanced tyrosine phosphorylation of cytoskeletally anchored proteins. *J BiolChem*, 273(9):5081-5085, 1998.
[8] Pelham, R.J., Jr., and Y. Wang. Cell locomotion and focal adhesions are regulated by substrate flexibility. *ProcNatlAcadSci U S A*, 94(25):13661-13665, 1997.
[9] Hynes R.O. Integrins: bidirectional, allosteric signaling machines.*Cell*,110(6):673–687, 2002.

[10] Shattil S.J., Kim C. and Ginsberg M.H. The final steps of integrin activation: the end game.*Nat Rev Mol Cell Biol*,11(4):288–300, 2010.
[11] Friedland J.C., Lee M.H. and Boettiger D. Mechanically activated integrin switch controls $\alpha 5\beta 1$ function.*Science*,323(5914):642–644, 2009.
[12] Gumbiner B.M. Regulation of cadherin-mediated adhesion in morphogenesis.*Nature Rev Mol Cell Biol*,6(8):622–634, 2005.
[13] Leckband D.E., le Duc Q., Wang N. and de Rooij J. Mechanotransduction at cadherin-mediated adhesions.*CurrOpin Cell Biol*,23(5):523–530, 2011.
[14] Yonemura S., Wada Y., Watanabe T., Nagafuchi A. and Shibata M. Alpha-catenin as a tension transducer that induces adherens junction development.*Nat Cell Biol*,12(6):533–542, 2010.
[15] Fey E.G., Wan K.M. and Penman S. Epithelial cytoskeletal framework and nuclear matrix-intermediate filament scaffold: three dimensional organization and protein composition. *J Cell Biol*,98(6):1973-1984, 1984.
[16] Rodriguez J.P., González, Rios S. and Cambiazo V. Cytoskeletal organization of human mesenchimal stem cells (MSC) changes during their osteogenic differentiation. *J Cell Biochem*, 93(4):721-731, 2004.
[17] Mammoto A. and Ingber D.E., Cytoskeletal control of growth and cell fate switching. *CurrOpin Cell Biol*, 21(6):864-870, 2009.
[18] Chen C.S., Mrksich M., Huang S., Withesides G.M. and Ingber D.E. Geometric control of cell life and death. *Science*, 276(5317):1425–1428, 1997.
[19] Ingber D.E. Tensegrity: the architectural basis of cellular mechanotrasduction. *Annu Rev Physiol*, 59:575–599, 1997.
[20] Giannone G. and Sheetz M.P. Substrate rigidity and force define form through tyrosine phosphatase and kinase pathways. *Trends Cell Biol*, 16(4):213-223, 2006.
[21] Fumarola L., Urani C. and Crosta G.F. Quantitative kinetics of damage and recovery of cytoskeletal structure by means of image analysis.*ToxicolIn Vitro*, 19(7):935–941, 2005.
[22] Chicurel M., Chen C.S. and Ingber, D.E. Cellular control lies in the balance of forces. *CurrOpin Cell Biol*,10(2):232-239, 1998.
[23] Chen C.S., Alonso J.L., Ostuni E., Whithesides G.M. and Ingber D.E. Cell shape provides global control of focal adhesion assembly. *Biochem Biophys Res Commun*, 307(2): 355-361, 2003.
[24] Le Beyec J., Xu R., Lee S.Y., Nelson C.M., Rizki A., Alcaraz J. and Bissell M.J. Cell shape regulates global histone acetylation in human mammary epithelial cells. *Exp Cell Res*, 313(14):3066-3075, 2007.
[25] Ingber D.E. Tensegrity-based mechanosensing from macro to micro.*ProgrBiophysMolBiol*, 97(2-3):163–179, 2008.
[26] Towbin B. D., Meister P. and Gasser S. M. The nuclear envelope-a scaffold for silencing?*CurrOpin Genet Dev*,19(2):180-186, 2009.
[27] Newman S.A., Forgas G. and Muller G.B. Before programs: The physical origination of multicellular forms. *Int J Dev Biol*, 50(2-3):289-299, 2006.
[28] Zhang H., Landmann F., Zahreddine H., Rodriguez D., Koch M. andLabouesse M. A tension-induced mechanotransduction pathway promotes epithelial morphogenesis. *Nature*, 471(7336):99-103, 2011.
[29] Bizzarri M., Pasqualato A., Cucina A. and Pasta V. Physical forces and non linear dynamics mould fractal cell shape.*Histol Histopathol*, 28(2):155-174, 2013.
[30] DeNardo D.G., Johansson M. and Coussens L.M. Immune cells as mediators of solid tumor metastasis. *Cancer Metastasis Rev*, 27(1):11–18, 2008.
[31] Whiteside T.L. The role of immune cells in the tumor microenvironment.*Cancer Treat Res*, 130:103–124, 2006.
[32] Paszek M.J., Zahir N., Johnson K.R., Lakins J.N., Rozenberg G.I., Gefen A., et al. Tensional homeostasis and the malignant phenotype. *Cancer Cell*, 8(3):241–254, 2005.

[33] Borghi N. and Nelson J.W. Intercellular adhesion in morphogenesis: molecular and biophysical considerations. *Curr Top Dev Biol*, 89:1–32, 2009.
[34] Ingber D.E. Can cancer be reversed by engineering the tumor microenvironment? *Semin CancerBiol*,18(5):356–364, 2008.
[35] Helmlinger G., Netti P.A., Lichtenbeld H.C., Melder R.J. and Jain R.K. Solid stress inhibits the growth of multicellular tumor spheroids. *Nat Biotechnol*, 15(8):778–783, 1997.
[36] Padera T.P., Stoll B.R., Tooredman J.B., Capen D., di Tomaso E. and Jain R.K. Pathology: cancer cells compress intratumour vessels. *Nature*, 427(6976):695, 2004.
[37] Levental K.R., Yu H., Kass L., Lakins J.N., Egeblad M., Erler J.T., Fong S.F., Csiszar K., Giaccia A.,Weninger W., Yamauchi M., Gasser D.L. and Weaver V.M. Matrix crosslinking forces tumorprogression by enhancing integrin signaling. *Cell*, 139(5):891–906, 2009.
[38] Chee P.N., Hinz B. and Swartz MA. Interstitial fluid flow induces myofibroblast differentiation andcollagen alignment in vitro. *J Cell Sci*, 118(Pt 20):4731–4739, 2005.
[39] Wyckoff J.B., Jones J.G., Condeelis J.S. and Segall J.E. A critical step in metastasis: in vivo analysis of intravasation at the primary tumor. *Cancer Res*, 60(9):2504–2511, 2000.
[40] Smith R.L., Donlon B.S., Gupta M.K., Mohtai M., Das P., Carter D.R., Cooke J., Gibbons G., Hutchinson N. andSchurman D.J. Effects of fluid-induced shear on articular chondrocyte morphology and metabolism in vitro.*J Orthop Res*,13(6):824-831, 1995.
[41] Kumar S. and Weaver V.M. Mechanics, malignancy, and metastasis: the force journey of a tumor cell.*Cancer Metastasis Rev*, 28(1–2):113–127, 2009.
[42] Kostic A., Lynch C.D. and Sheetz M.P. Differential matrix rigidity response in breast cancer cell linescorrelates with the tissue tropism. *PLoS One*, 4(7):e6361, 2009.
[43] Pathi S.P., Kowalczewski C., Tadipatri R. and Fischbach C. A novel 3-D mineralized tumor model tostudy breast cancer bone metastasis.*PLoS One*, 5(1):e8849, 2010.
[44] Berenblum I. and Shubik P. An experimental study of the initiating state of carcinogenesis and a re-examination of the somatic cell mutation theory of cancer. *Br JCancer*,3(1):109–118, 1949.
[45] Bissell M.J., Hatie C. and Calvin M. Is the product of the src gene a promoter? *ProcNatlAcadSci USA*,76(1):348–352, 1979.
[46] Barcellos-Hoff M.H. Stromal mediation of radiation carcinogenesis.*J Mammary Gland Biol Neoplasia*,15(4):381-387, 2010.
[47] Halachmi E. and Witz I.P. Differential tumorigenicity of 3T3 cells transformed in vitro with polyoma virus and in vivo selection for high tumorigenicity. *Cancer Res*, 49(9):2383–2389, 1989.
[48] Vogelstein B. and Kinzler K.W. Cancer genes and the pathways they control. *Nat Med*,10(8):789–799, 2004.
[49] Gupta G.P. and Massague J. Cancer metastasis: building a framework. *Cell*, 127(4):679–695, 2006.
[50] Paget S. The distribution of secondary growths in cancer of the breast. 1889.*Cancer Metastasis Rev*,8(2):98-101,1989.
[51] Pierce B.G. and Wallace C. Differentiation of malignant to benign cells.*Cancer Res*, 31(2):127-134, 1971.
[52] D'Anselmi F., Masiello M.G., Cucina A., Proietti S., Dinicola S., Pasqualato A., Ricci G., Dobrowolny G., Catizone A., Palombo A. and Bizzarri M. Microenvironment promotes tumor cell reprogramming in human breast cancer cell lines.*PLoS One*,8(12):e83770, 2013.
[53] Hendrix M.J.C., Seftor E.A., Seftor R.E.B., Kasemeier-Kulesa J., Kulesa P.M. and Postovit L.M. Reprogramming metastatic tumour cells with embryonic microenvironments. *Nat Rev Cancer*, 7(4):246-255, 2007.
[54] Bizzarri M., Cucina A., Biava P.M., Proietti S., D'Anselmi F., Dinicola S. and Pasqualato A. Embryonic morphogenetic field induces phenotypic reversion in cancer cells. Review article.*Curr Pharm Biotechnol*,12(2):243-253, 2011.

[55] Bissell M.J., Kenny P.A. and Radisky D.C. Microenvironmental regulators of tissue structure and function also regulate tumor induction and progression: the role of extracellular matrix and its degrading enzymes.*Cold Spring HarbSymp Quant Biol*, 70:343-356, 2005.
[56] Kenny P.A. and Bissell M.J. Tumor reversion: correction of malignant behavior by microenvironmental cues. *Int J Cancer*, 107(5):688–695, 2003.
[57] Krause S., Maffini M.V., Soto A.M. andSonnenschein C. The microenvironment determines the breast cancer cells' phenotype: organization of MCF7 cells in 3D cultures. *BMC Cancer*,10:263, 2010.
[58] Witz I.P., Hanna M.G. Jr (eds) Contemp Top Immunobiol, 10. In situ expression of tumor immunity. Plenum, New York, 1980.
[59] Folkman J., Merler E., Abernathy C et al. Isolation of a tumor factor responsible for angiogenesis. *J Exp Med*, 133(2):275–288, 1971.
[60] Witz I.P. The Tumor Microenvironment: The Making of a Paradigm.*Cancer Microenviron*, 2 (Suppl 1):9–17, 2009.
[61] Bissell M.J. andBarcellos-Hoff M.H. The influence of extracellular matrix on gene expression: is structure the message? *J Cell SciSuppl*,8:327–343, 1987.
[62] Van den Hooff A. Stromal involvement in malignant growth. *Adv Cancer Res*, 50:159–196, 1988.
[63] Rana B., Mischoulon D., Xie Y., Bucher N.L.andFarmer S.R. Cell-extracellular matrix interactions can regulate the switch between growth and differentiation in rat hepatocytes: reciprocal expression of C/EBP alpha and immediate-early growth response transcription factors. *Mol Cell Biol*, 14(9):5858-5869, 1994.
[64] Huang S., Chen C.S. and Ingber D.E. Control of cyclin D1, $p27^{Kip1}$, and cell cycle progression in human capillary endothelial cells by cell shape and cytoskeletal tension. *MolBiol Cell*, 9(11):3179-3193,1998.
[65] Dike L.E. and Ingber D.E. Integrin-dependent induction of early growth response genes in capillary endothelial cells. *J Cell Sci*, 109(Pt 12):2855-2863,1996.
[66] Roskelley C.D., Desprez P.Y. and Bissell M.J. Extracellular matrix-dependent tissue-specific gene expression in mammary epithelial cells requires both physical and biochemical signal transduction. *ProcNatlAcadSci USA*, 91(26):12378–12382, 1994.
[67] Boudreau N., Myers C. and Bissell M.J. From laminin to lamin: regulation of tissue-specific gene expression by the ECM. *Trends Cell Biol*, 5(1):1–4, 1995.
[68] Chen C.S., Mrksich M., Huang S., Withesides G.M. and Ingber D.E. Geometric control of cell life and death. *Science*,276(5317):1425–1428, 1997.
[69] Paszek M.J., Zahir N., Johnson K.R., Lakins J.N., Rozenberg G.I., Gefen A. et al. Tensional homeostasis and the malignant phenotype. *Cancer Cell*, 8(3):241–254, 2005.
[70] Meredith J.E. Jr, Fazeli B. andSchwartz M.A. The extracellular matrix as a cell survival factor.*MolBiol Cell*, 4(9):953-961, 1993.
[71] Boudreau N., Sympson C.J., Werb Z. and Bissell M.J. Suppression of ICE and apoptosis in mammary epithelial cells by extracellular matrix. *Science*,267(5199):891–893, 1995.
[72] Colpaert, C.G., Vermeulen, P.B., Fox, S.B., Harris, A.L., Dirix, L.Y., and Van Marck, E.A. The presence of a fibrotic focus in invasive breast carcinoma correlates with the expression of carbonic anhydrase IX and is a marker of hypoxia and poor prognosis. *Breast Cancer Res Treat*, 81(2):137–147, 2003.
[73] Martin L.J. and Boyd N.F. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. *Breast Cancer Res*,10(1):201, 2008.
[74] Szauter K.M., Cao T., Boyd C.D. and Csiszar K. Lysyl oxidase in development, aging and pathologies of the skin. *PatholBiol (Paris),* 53(7):448–456, 2005.

[75] Balkwill F., Charles K.A. and Mantovani A. Smoldering and polarized inflammation in the initiation and promotion of malignant disease. *Cancer Cell*, 7(3):211–217, 2005.
[76] Coussens L.M., Raymond W.W., Bergers G., Laig-Webster M., Behrendtsen O., Werb Z., Caughey, G.H. and Hanahan D. Inflammatory mast cells up-regulate angiogenesis during squamous epithelial carcinogenesis. *Genes Dev*, 13(11):1382–1397, 1999.
[77] Fambrough D., McClure K., Kazlauskas A. and Lander E.S. Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell*, 97(6):727–741, 1999.
[78] Wernig M., Meissner A., Foreman R., Brambrink T., Ku M., Hochedlinger K. et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*,448(7151):318–324, 2007.
[79] Qu K., AbiHaidar A., Fan J., Ensman L., Tuncay K., Jolly M. andOrtoleva P. Cancer onset and progression: a genome-wide, nonlinear dynamical systems perspective on onconetworks. *J TheorBiol*,246(2):234-244, 2007.
[80] Kashiwagi A., Urabe I., Kaneko K. and Yomo T. Adaptive Response of a Gene Network to Environmental Changes by Fitness-Induced Attractor Selection. *PLoS ONE*, 1:e49, 2006.
[81] Levin M. and Stevenson C.G. Regulation of cell behavior and tissue patterning by bioelectrical signals: challenges and opportunities for biomedical engineering.*Annu Rev Biomed Eng*, 14:295-323, 2012.
[82] Bizzarri M., Cucina A., Palombo A. and Masiello M.G. Gravity sensing by cells: mechanisms and theoretical grounds. *Rend FisAccLincei*, vol. 25 no. 1, pp. S29–S38, 2014.
[83] Levin M. Reprogramming cells and tissue patterning via bioelectrical pathways: molecular mechanisms and biomedical opportunities.*Wiley Interdiscip Rev SystBiol Med*, 5(6):657-676, 2013.
[84] Levin M. Morphogenetic fields in embryogenesis, regeneration and cancer: non-local control of complex patterning. *Biosystems*,109(3):243-261, 2012.
[85] Tee S-Y., Fu J., Chen C.S. and Janmey P.A. Cell shape and substrate rigidity both regulate cell stiffness. *Biophys J*, 100(5):(L25-27), 2011.
[86] Chicurel M.E., Chen C.S. andIngber D.E. Cellular control lies in the balance of forces. *CurrOpin Cell Biol*,10(2):232-239, 1998.
[87] Derynck R., Akhurst R.J. and Balmain A. TGF-β signaling in tumor suppression and cancer progression. *Nat Genet*, 29(2):117–129, 2001.
[88] Moses H.L., Yang E.Y. and Pietenpol J.A. TGF-β stimulation and inhibition of cell proliferation: new mechanistic insights. *Cell*, 63(2):245–247, 1990.
[89] Deckers M., van Dinther M., Buijs J. et al. The tumor suppressor Smad4 is required for transforming growth factor h-induced epithelial to mesenchymal transition and bone metastasis of breast cancer cells. *Cancer Res*, 66(4):2202–2209, 2006.
[90] Bierie B. and Moses H.L. Tumour microenvironment: TGFh: the molecular Jekyll and Hyde of cancer. *Nat RevCancer*, 6(7):506–520, 2006.
[91] Grinnell F. and Ho C.H. Transforming growth factor beta stimulates fibroblast-collagen matrix contraction by different mechanisms in mechanically loaded and unloaded matrices. *Exp Cell Res*,273(2):248-255, 2002.
[92] Fedorova E. and Zink D. Nuclear architecture and gene regulation. *BiochimBiophysActa*,1783(11):2174-2184, 2008.
[93] Mammoto A. andIngber D.E. Cytoskeletal control of growth and cell fate switching.*CurrOpin Cell Biol*, 21(6):864-870, 2009.
[94] Feng J., Tang Y., Xu Y., Sun Q., Liao F. andHan D. Substrate stiffness influences the outcome of antitumor drug screening in vitro. *ClinHemorheolMicrocirc*, 55(1):121-131, 2013.
[95] Mintz B. and Illmensee K. Normal genetically mosaic mice produced from malignant teratocarcinoma cells. *ProcNatlAcadSci USA*, 72(9):3585–3589, 1975.

[96] Tschumperlin D.J., Dai G., Maly I.V., Kikuchi T., Laiho L.H., McVittie A.K., Haley K.J., Lilly C.M., So P.T., Lauffenburger D.A. et al. Mechanotransduction through growth-factor shedding into the extracellular space. *Nature*, 429(6987):83–86, 2004.
[97] Wipff P.J., Rifkin D.B., Meister J.J. and Hinz B. Myofibroblast contraction activates latent TGF-beta1 from the extracellular matrix. *J Cell Biol*, 179(6):1311–1323, 2007.
[98] Newman S.A., Forgas G. and Muller G.B. Before programs: the physical origination of multicellular forms. *Int J Dev Biol*,50(2-3):289–299, 2006.
[99] Budhu α. and Wang X.W. Transforming the Microenvironment: A Trick of the Metastatic Cancer Cell . *Cancer Cell*, 22(3):279-280, 2012.
[100] Guadagno T.M. and Assoian R.K. G1/S control of anchorage-independent growth in the fibroblast cell cycle. *J Cell Biol*, 115(5):1419–1425, 1991.
[101] Stoker M.G. and Rubin H. Density dependent inhibition of cell growth in culture. *Nature*, 215(5097):171–172, 1967.
[102] Testa F., Palombo A., Dinicola S., D'Anselmi F., Proietti S. et al.Fractal analysis of shape changes in murine osteoblasts cultured under simulated microgravity. *Rend FisAccLincei*, vol. 25, n.1, pp. S39–S47, 2014.
[103] Bajenaru M.L., Zhu Y., Hedrick N.M., Donahoe J., Parada L.F. and Gutmann D.H. Astrocyte-specific inactivation of the neurofibromatosis 1 gene (Nf1) is insufficient for astrocytoma formation. *Mol Cell Biol*, 22(14):5100–5113, 2002.
[104] Daginakatte G.C.andGutmann D.H.. Neurofibromatosis-1 (Nf1) heterozygous brain microglia elaborate paracrine factors that promote Nf1-deficient astrocyte and glioma growth. *Hum Mol Genet*,16(9):1098-1112, 2007.
[105] Soto A.M., Maffini M.V. and Sonnenschein C.Neoplasia as development gone awry: the role of endocrine disruptors.*Int J Androl*,31(2):288-293, 2008.
[106] Michor F., Liphardt J., Ferrari M. and Widom J. What does physics have to do with cancer? *Nat Rev Cancer*,11(9):657-670, 2011.
[107] Krause S., Maffini M.V., Soto A.M.andSonnenschein C. The microenvironment determines the breast cancer cells' phenotype: organization of MCF7 cells in 3D cultures. *BMC Cancer*,10:263, 2010.
[108] Bizzarri M., Palombo A. and Cucina A. Theoretical aspects of Systems Biology.*ProgBiophysMolBiol*,112(1-2):33-43, 2013.
[109] Kuhn T.S. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.

Poster Session

Biophysics & Bioengineering Letters, 2013, (6), 1-9 .

NMR protein conformers described by network parameters: a preliminary study.

Lisa Beatrice Caruso, Alessandro Giuliani, Cesare Manetti, Alfredo Colosimo

Abstract

NMR spectroscopy is one of the techniques of choice for the determination of protein structures. Its use has a number of positive aspects, among which the possibility to observe the influence of the solvent on the molecular structure, as well as the local movement of small molecular domains. However, due to the intrinsic flexibility of protein tertiary structures in solution, the NMR information does not lead to a single structure but to a set of conformers. Using the topological representation of such conformers we analyzed the corresponding network parameters, to enlight their association with some specific molecular feature. In this frame we showed that: i) the node degree parameter positively correlates with molecular 'compactness', ii) the average shortest path length parameter positively correlates with molecular flexibility, and iii) as expected, the two parameters are anticorrelated between each other.

Topological Characterization of NMR conformers

Lisa Beatrice Caruso[a], Cesare Manetti[b], Alfredo Colosimo[a]
[a] Dept. SAIMLAL – [b]Dept. of Chemistry (Sapienza University, Rome)

This work deals with the use of the topological networks in the characterization of the Myoglobin conformers obtained by NMR spectroscopy.

Topological Description of Proteins

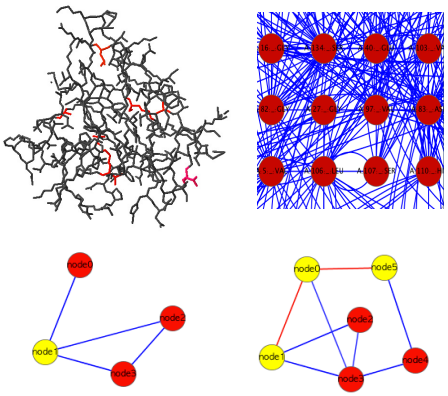
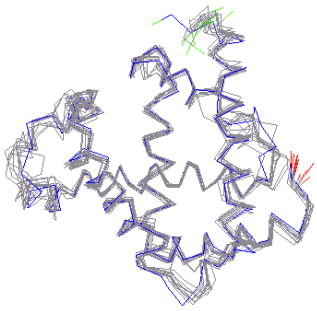


Figure1: *upper pannel the difference in* geometrical and topological representation of a protein structure is sketched.*Bottom pannel:* graphical view of the two parameters used: *left* node degree (ND), the number of adjacent edges to a node (node1=3), *right* shortest path L(n,m) the path with smallest number of links between two nodes L(1,5)=2. The Average Shortest Path Length (ASPL*) is the* Average length of shortest paths between all the node pairs.

SWMB(CO) NMR conformers described by topological parameters.



Conformer	*# Nodes*	*# Edges*	*ASPL*	*Average ND*
1	149	800	3.39	10.37
2	151	832	3.34	10.80
3	151	808	3.46	10.73
4	141	808	3.46	10.52
5	153	822	3.47	10.75
6	152	807	3.43	11.05
7	149	789	3.42	10.31
8	152	821	3.40	10.58
9	150	804	3.42	10.39
10	151	799	3.44	10.41
11	151	797	3.48	10.09
12	149	798	3.28	10.27

Fiigure 2: *right* NMR structures in solution of carbomonoxy myoglobin (SWMb(CO). *Left:* Network parameters of the 12 NMR conformers. Columns 4 and 5 contain, respectively, the mean values of Node Degree (ND) and Average Shortest Path Length (ASPL) for all nodes of the corresponding conformer.

SWMb(CO) NMR conformers: Distribution of topological parameters along the sequence.

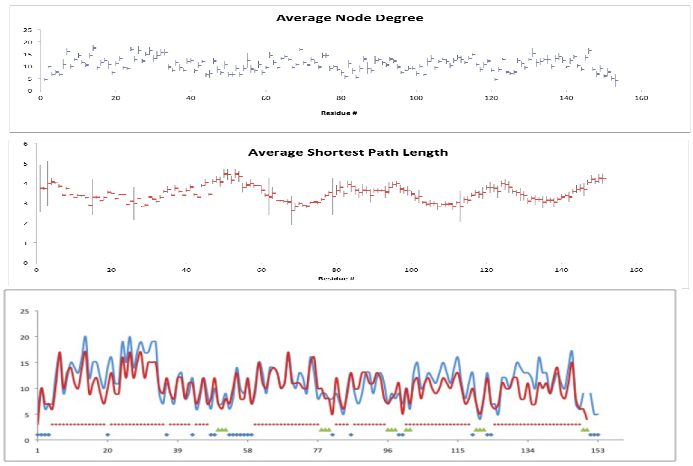


Figure 3: *upper and middle panels*: Profile of the ND and ASPL along the SW Mb(CO) primary structure. The parameter value averaged over the 12 conformers and the related standard deviation, are indicated for each node. Notice the striking negative correlation between the two parameters distribution along the sequence (R= -0.62). The *bottom panel* shows the differences in the ND profile between #6 and #11 conformers . In the bottom part of the graph red, green and blue dots indicate the location of the secondary structure elements, i.e. *α* elix, *β* sheet and Random,

High resolution of topological parameters to proteins' dynamic features

ASPL	*40,43*
NODE DEGREE	*33,4*
RMSD	*8,4*

Figure4: the variability of both topological parameters, expressed by Percent Variation Coefficient (PVC=(SD/MEAN)X100), is calculated in all cases taking into account each of the 121 couples of conformers, The significantly higher values of the ND and ASPL parameters with respect to the corresponding geometrical parameter (RMSD), could be taken as an indication of the higher sensitivity to the protein dynamic features.

Conclusion: Our first aim was to generalize the results by deriving from differences in topological features some indication concerning the structural conformers suggested by NMR data. The final aim remains to associate the above information to a time-dependent sequential ordering directly reflecting a dynamic behaviour .

[1] Barabàsi A.L. and Oltvai Z.N. Network biology: understanding the cell's functional organization. Nat Rev Genet, 5:101-113, 2004.
[2] K. W uthrich. Nmr studies of structure and function of biological macromolecules (nobel lecture). Angew.Chem.Int.Ed.42, 3340-3363, 2003.
[3] del Sol A., Fujihashi H., Amoros D., and Nussinov R.. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. Molecular Systems Biology, page 112, 2006.
[4] Caruso L.B., Manetti C., Giuliani A. and Colosimo A. NMR protein conformers described by network parameters:a preliminary study. Biophysics and Bioengineering Letters, Vol.6, 2013.

Proceedings of the Royal Society B, 2009, (276), 833–841

Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations

Øivind Andersen, Ola Frang Wetten, Maria Cristina De Rosa, Carl Andre, Cristiana Carelli Alinovi, Mauro Colafranceschi, Ole Brix and Alfredo Colosimo

Abstract

A major challenge in evolutionary biology is to identify the genes underlying adaptation. The oxygen-transporting haemoglobins directly link external conditions with metabolic needs and therefore represent a unique system for studying environmental effects on molecular evolution. We have discovered two haemoglobin polymorphisms in Atlantic cod populations inhabiting varying temperature and oxygen regimes in the North Atlantic. Three-dimensional modelling of the tetrameric haemoglobin structure demonstrated that the two aminoacid replacements Met55β1 Val and Lys62β1 Ala are located at crucial positions of the α1β1 subunit interface and haem pocket, respectively. The replacements are proposed to affect the oxygen-binding properties by modifying the haemoglobin quaternary structure and electrostatic feature. Intriguingly, the same molecular mechanism for facilitating oxygen binding is found in avian species adapted to high altitudes, illustrating convergent evolution in water- and air-breathing vertebrates to reduction in environmental oxygen availability. Cod populations inhabiting the cold Arctic waters and the low-oxygen Baltic Sea seem well adapted to these conditions by possessing the high oxygen affinity Val55–Ala62 haplotype, while the temperature-insensitive Met55–Lys62 haplotype predominates in the southern populations. The distinct distributions of the functionally different haemoglobin variants indicate that the present biogeography of this ecologically and economically important species might be seriously affected by global warming.

Hemoglobin Polymorphisms affect the Oxygen Binding Properties in Atlantic Cod populations

Øivind Andersen[1 2], Ola Frang Wetten[2 3], Maria Cristina De Rosa[4], Carl Andre[5], Cristiana Carelli Alinovi[6], Mauro Colafranceschi[7], Ole Brix[8 9], Alfredo Colosimo[7]

[1]AKVAFORSK, Aas, Norway; [2]University of Life Sciences, Norway; [3]Hedmark University College, Hamar, Norway; [4]CNR, Rome, Italy; [5]Göteborg University, Strömstad, Sweden; [6]Catholic University of Rome, Italy; [7]CISB Rome, Italy; [8]University of Bergen, Norway; [9]The Michelsen Centre for Industrial Measurement Science and Technology, Bergen, Norway.

Background and methods

Elucidating the molecular mechanisms of adaptation is a central aim in evolutionary biology. Temperature change affects the haemoglobin-oxygen-binding affinity as well as the ambient oxygen availability. This environmental factor sets the limits for life by influencing the distribution and production of marine species (Brandner 2007). The **Atlantic cod** (***Gadus morhua***) is widely distributed in the North Atlantic, encountering variable temperature and oxygen regimes.

•We present two polymorphisms in the cod hemoglobins, which show profound functional differences in Northeast Atlantic and Baltic populations.

•The fish were genotyped using erythrocyte cDNA and genomic DNA as templates for the PCR amplification and direct sequencing of polymorphic sites identified in the Hb-**β**1, Hb-**β**3 and Hb-**α**2 genes (figure 1).

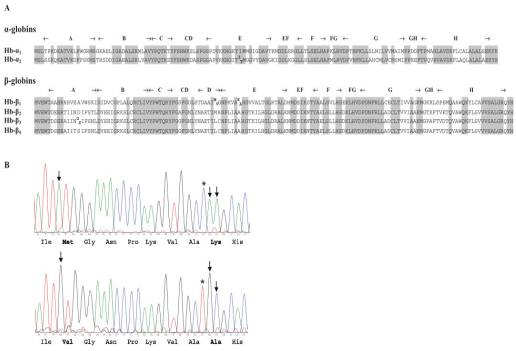


Figure 1. Amino acid polymorphisms in globin subunits of Atlantic cod. **A.** Deduced amino acid sequences of the six isolated globin genes, including the alternative amino acids at the polymorphic sites. The helices are conventionally designated in bold letters. Highly conserved positions are shaded. **B.** Sequence chromatogram of the mutated region of *Hb-**β**1* encoding the non-recombinant allelic globin variants Met55-Lys62 (upper) and Val55-Ala62 (lower). Non-synonymous mutation (↓), synonymous mutation (*).

Results 1

By PCR amplification of multiple cod globin genes we identified two amino acid replacements Met55Val and Lys62Ala in the **β**1 globin chain:

•ambiguous genotype-phenotype relationships were found by analysing erythrocyte globin DNA, mRNA and proteins.

•Three-dimentional modelling of the tetrameric structure demonstrated that these amino acids are located at crucial positions of the **α**1**β**1 subunit interface (figure 2) and haem pocket (figure 3), respectively. Hence, the replacements affect the oxygen binding properties by modifying the hemoglobin quaternary structure and electrostatic features.

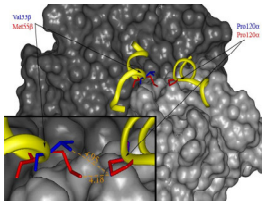


Figure 2. Met55**β**Val substitution in the **α**1**β**1 subunit contact of Atlantic cod hemoglobin. The 3D model shows the superimposition of Met55b (red) and Val55b (blue) at the interface of the **α**1 (black) and **β**1 (gray) subunits. The D helix of **β**1 and the GH corner and initial H helix of **α**1 chain are shown in ribbon representation (yellow). The frame depicts the distances of the CG atom of Pro120**α** to the CE atom of Met55**β** and to the CG2 atom of Val55**β**.

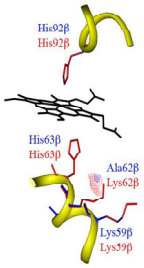


Figure 3. Lys62**β**Ala substitution and water interaction in the heme pocket. The figure shows the superimposition of GRID contour maps at -11.0 kcal/mol for the water probe calculated for the distal heme pocket containing Lys62**β** (red) or Ala62**β** (blue). His63b, Lys59b and Lys62b were identified as the key residues contributing to the hydrophilic contour maps.

Results 2

Cod populations inhabiting the cold Arctic waters and the low-oxygen Baltic Sea seem well adapted to these conditions by possessing the high oxygen affinity Val55–Ala62 haplotype, while the temperature-insensitive Met55–Lys62 haplotype predominates in the southern populations (figure 4).

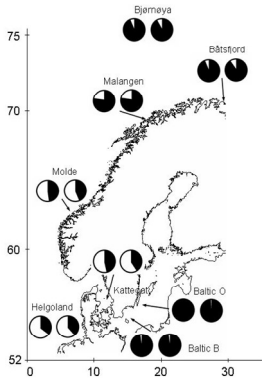


Figure 4. Relative allele frequencies of the Met55Val (left pie) and Lys62Ala (right pie) polymorphisms in the cod populations examined. The Met55 and Lys62 alleles are depicted in white, whereas the Val55 and Ala62 alleles are shown in black. Altitudes (ordinates) and longitudes (abscissas) are indicated. Baltic Ö (Öland), Baltic B (Bornholm).

Discussion and conclusions

- Amino acid replacements Met55**β**1Val and Lys62**β**1Ala are proposed to affect the oxygen-binding properties of Atlantic cod haemoglobin.
- Cod populations inhabiting cold or low-oxygen waters were adapted to these conditions by evolving a high oxygen affinity haplotype.
- The same molecular mechanism for facilitating oxygen binding is found in avian species adapted to high altitudes (Liang *et al.* 2001), illustrating convergent evolution and adaptation of water- and air-breathing vertebrates to reduction in environmental oxygen availability.
- The distinct distributions of the functionally different haemoglobin variants indicate that the present biogeography of this ecologically and economically important species might be seriously affected by global warming.

References

1. Brander KM (2007) Global fish production and climate change. Proc Natl Acad Sci USA 104: 19709-19714.
2. Liang Y., Hua Z., Liang X., Xu Q. & Lu G. (2001) The crystal structure of bar-headed goose hemoglobin in deoxy form: the allosteric mechanism of a hemoglobin species with high oxygen affinity. J. Mol. Biol. 313, 123–137.

Biophysics & Bioengineering Letters, 2014, (7), 1-8 .

Associating brain topological networks to cognitive performance.

Fabrizio Parente, Alfredo Colosimo

Abstract

Recent applications of graph theory to brain networks showed the possibility of a relation between network topological indexes and cognitive abilities. In this paper we want to study the relation between the topology of brain networks and a known parameter of executive functions, such as perseveration, using the WCST. To this end, nine healthy subjects were subjected to a fMRI acquisitions with a 3 Tesla Siemens scanner under condition of resting state and evaluated with the WCST. The images were analyzed using the following Matlab toolboxes: SPM8 and Functional Connectivity Toolbox. FromWCST data the indexes of perseverative, nonperseverative errors and perseverative response were calculated. A small-world feature appears in the cost range (T) 0.45 - 0.50. In this interval, the index shows a positive correlation with perseverative responses (T=0.50, r=0.864, p=0.006). Moreover, among the values of single cerebral regions (T=0.50), the middle part of orbital frontal gyrus left (r=0.920, p=0.001) show a significant trend for positive correlation with perseverative responses. These results suggest a relation between network's segregation and perseveration; more specifically a greater segregation of subnetworks is related to a lower adaptability of behavior to the environment changes. Our findings can provide hints to understand the pathological alterations of mental disease related to impairment of executive functions.

Associating brain topological networks to cognitive performance.

Fabrizio Parente, Alfredo Colosimo
Dept. SAIMLAL. Sapienza, University of Rome
e-mail: fabrizio.parente86@gmail.com

In task-related fMRI investigations, the performance in cognitive tests is thought to influence the activation of one or more specific cerebral areas. With different analytical techniques it is possible to integrate these cerebral areas into a functional networks. As a matter of fact, BOLD signal fluctuations of these areas have been correlated between each other in the resting state [1] and such correlations: i) remain stable even in the lack of given stimuli, and ii) involve a relevant fraction of the brain metabolism. In this contest, the graph theory provides a theoretical support in the investigation of complex network topologies [2] allowing to answer critical questions like: 1) How is the topology of brain networks organized? 2) How efficiently the brain integrates the information coming from subnetworks? Recent applications of this technique showed the possible existence of correlation between network topological indexes and cognitive abilities [3-5].

In this contribution we study the relation between the topology of brain networks and a known parameter of executive functions, such as perseveration, using the Wisconsin Card Sorting Task (WCST). To this end, nine healthy subjects were subjected to a fMRI acquisitions with a 3 Tesla Siemens scanner under condition of resting state and evaluated by the WCST. The fMRI images were analyzed using the SPM8, Functional Connectivity and Brain Connectivity Toolbox. In particular, from the correlation matrix (**Figure 1**) of each subject the corresponding binary network was obtained and the following indexes were calculated (**Figure 2**): Clustering Coefficient (C), Characteristic Path Length (L), Gamma (γ) and Lambda (λ). Finally, from WCST data the indexes of perseverative and nonperseverative errors, as well as, of perseverative response, were calculated (**Figure 3**).

The partial correlations (controlling variable = age) show a positive trend between the gamma index and perseverative responses ($r = 0.864$, $p = 0.006$; **Figure 4**). Moreover, among the values of single cerebral regions, the orbital part of left middle frontal gyrus ($r = 0.920$, $p = 0.001$) shows a significant trend for positive correlation with perseverative responses.

These results suggest a relation between network's segregation and perseveration. More specifically, a greater segregation of subnetworks (high value of gamma index) seems related to a lower adaptability of behavior to the environmental changes (high degree of perseveration). Our findings can provide hints to understand the pathological alterations of mental diseases related to impairment of executive functions.

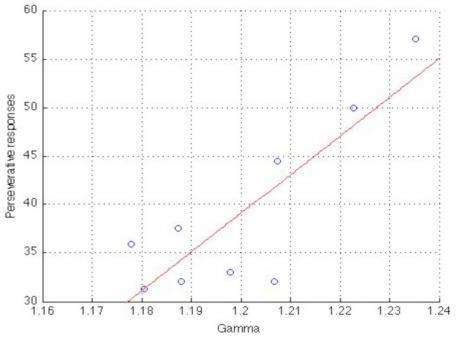


Figure 4. Correlation between gamma values of the individual network parameters and perseverative responses of the 9 healthy subjects analyzed in this work.

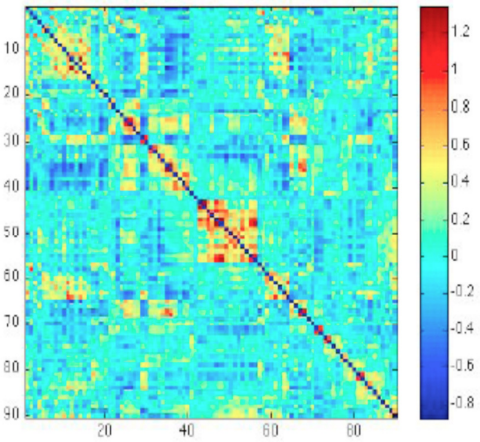


Figure 1. Pearson Correlation Matrix between ROIs of a single subject from fMRI (time series of BOLD) signals. The data are plotted in a normalized false color scale.

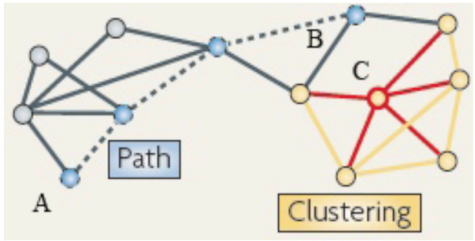


Figure 2. Calculation of the Topological Network parameters (image modified from [6]). The shortest path length is the minimal number of links connecting two nodes. In the picture the shortest path between A and B nodes (the dotted line) is 3 and the average of this index is L. The clustering coefficient (C) is the fraction of neighbor nodes that are neighbors to each other. In the picture the yellow nodes are the neighbors of the C node; the clustering coefficient of this is 0.5. The gamma and lambda indices are the C and L normalized to the corresponding random network.

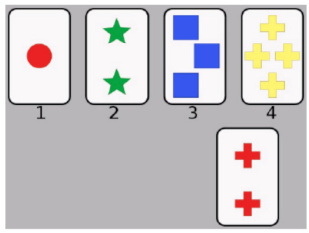


Figure 3. WCST carried out on the PEBL platform (http://pebl.sf.net). Subjects must associate one of four key cards (top row) to the response card (bottom) on the basis of color or shape or number. The bottom card is one of 128 responses cards.

[1]M.E. Raichle, et all., PNAS 14(2010) 180–190.
[2]E. Bullmore and O. Sporns, Nat Rev Neurosci. 10(2009) 186–98.
[3]M.P. Van den Heuvel, C.J. Stam, R.S. Kahn, and H.E. Hulshoff Pol, J. Neurosci. 29(2009) 7619–7624.
[4]A.A. Stevens, S.C. Tappon, A. Garg and D.A. Fair, PLoS One 7(2012):e30468.
[5]M. Song, Y. Zhou , J. Li, Y. Liu, L. Tian, C. Yu and T. Jiang, Neuroimage 41(2008) 1168–1176.
[6]E. Bullmore, O. Sporn, Nat Rev Neurosci. 10(2009):186-98.

Biophysics & Bioengineering Letters, 2013, (6), 9-17 .

Simulation of neural networks: a preliminary report on two computational strategies.

Daiana Simone, Alfredo Colosimo

Abstract

The basic assumptions of the present contribution are the following: i) a satisfactory mechanistic knowledge of the higher human cognitive abilities will be reached after understanding the concerted and cooperative behavior of brain regions usually identified by peculiar morphofunctional features; ii) it is really hard to overestimate the importance of computer-aided modeling and simulation in the study of the global and local network(s) connecting the above regions; iii) the related computational problems are better faced by exploiting different software tools, each endowed with excellent performance in specific problems, than focussing on a single programming environment. In this frame, we tested the steepness of the learning curve of two popular neural simulation environments, namely NEURON and BRIAN, in modeling the following cases: a) a selfsustained, reciprocal activation of a small number of ring-chained neurons, and b) the spiking activity of a small network possibly endowed with a random set of connections. We confirm that the NEURON and the BRIAN environments appear well suited in dealing with case a) and b), respectively, and, in particular, in studying the biophysical / morphological features of single neural cells or in exploring the functional features even of complicated or large-size network topologies.

Simulation of neural networks: a preliminary report

Daiana Simone and Alfredo Colosimo
Dept. SAIMLAL – Sapienza University, Rome

This work deals with the influence of random connections in modeling the efficiency of neural networks by simple reverberating circuits.

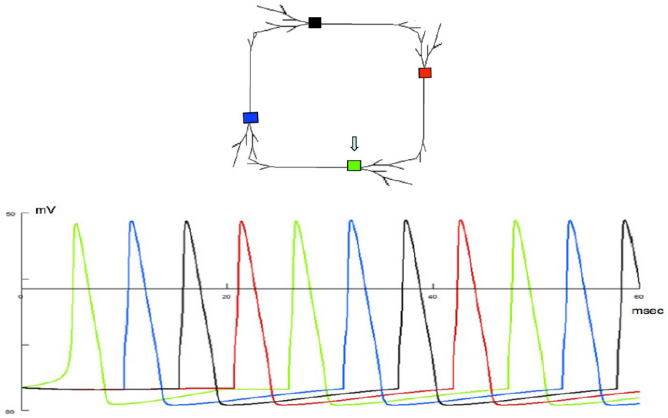


Figure 1. Reverberating activity of a linear closed-loop network.
•Top : A reverberating circuit obtained by the software tool in [1], as an oversimplified model of working memory [2].
•Bottom: notice the stable and regular pattern of spikes following the initial short (5msec) stimulation of a node (indicated by the arrow).
The essential feature of the *reverberating circuit* hypothesis, namely the emergence of a long-term, self-sustaining activity, is fully respected.

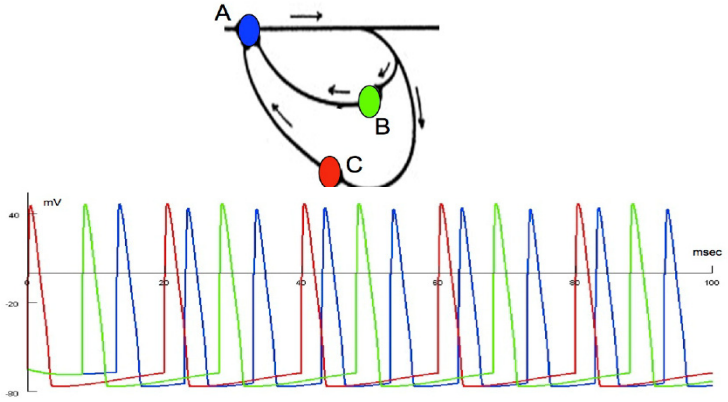


Figure 2. Multiple frequencies of a non-linear network.
• Top: A minimal circuit generator of multiple frequencies. Neurons A, B, C due to their specific patterns of links, produce specific Action potentials (APs)
•Bottom: The APs produced by neurons in the top panel can be identified by their color.
Notice that the self-sustaining activity is independent of the number of nodes, and directly influenced by the connection topology.

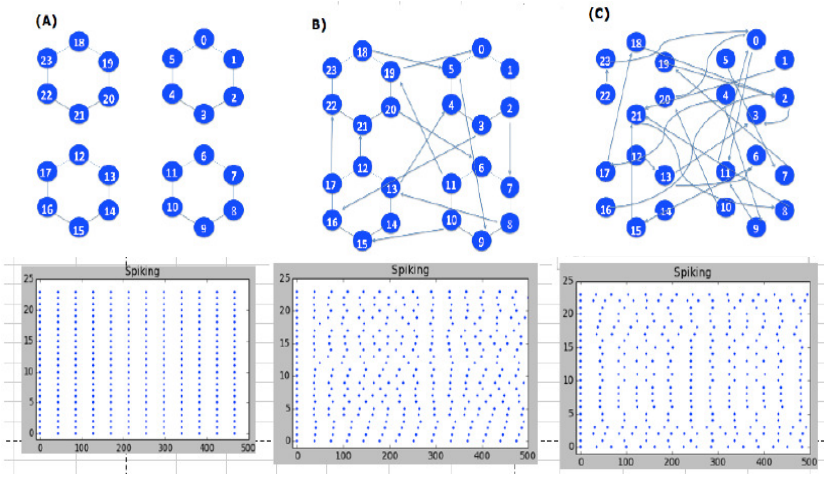


Figure 3. Firing of single neurons in ordered and disordered link topologies.
Top: An ordered (A), a quasi ordered (B) and a random (C) link topology were set up over the same node arrangement and studied by the BRIAN software [3] .
Bottom: The spiking activity associated to the three above link topologies are reported in the form of raster plots: Xaxis = time (msec); Yaxis = Firing of each of the 24 nodes included in the network.
Thus, the fully disconnected four subnetworks in (A) appear the least efficient in terms of number of spikes generated in the available time window of 500 msec (see also table 1)

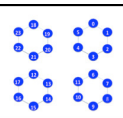
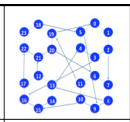
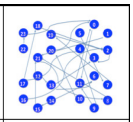
			
CHARACTERISTIC PATH LENGHT (CPL)	3	4,93	2.44
AVERAGE NODE DEGREE (ND)	2	3	1.92

Table 1. Topological parameters of the networks in Fig. 3

In the (A), (B) and (C) topologies in Fig. 3, the Average Node Degree (ND) and Characteristic Path Length (CPL) values confirm that in spite of the lower average number of links per node (ND), a fully random link distribution (C) may provide a shorter average (hence a more efficient) communication pathway (CPL) among nodes.

Conclusions: Connections topology appears as the most important factor in reproducing by simple circuits a working memory performance.

[1] Carnevale N T and Hines M. The NEURON Book. Cambridge, 2004.
[2] Klein R M. The hebb legacy. Canadian Journal of Experimental Psychology, 53(1): 13, 2006.
[3] Goodman D and Brette R. Brian: a simulator for spiking neural networks in python. Front. Neuroinform., 2:5. doi:10.3389/neuro.11.005., 2008.
[4] Strogatz S H. Exploring complex networks. Nature, 410:268-277, 2001.

"Biosignals 2010" Third intern. Meeting, Valencia, 2013

EEG SIGNALS IN EPILEPSY AND MIGRAINE
Analysis and Simulations by Multi-Agent Systems

Alessandro Vigano', Neri Accornero, Alfredo Colosimo

ABSTRACT

It has been suggested some time ago that epilepsy and migraine are correlated, indicating that a crucial role in both pathologies is played by an abnormal synchronization of the involved neuronal populations. This has been recently reassessed on the basis of accurate epidemiological data). Thus, recruiting a larger and larger number of phase-coupled neurons, should account for: i) the peculiar activity bursts appearing in EEG signals; ii) the close temporal correlation of the activity bursts with macroscopic clinical symptoms like epileptic seizures or individual perceptions like visual aura; iii) the typical rythmic occurrence and spatial patterns of the activity waves.

Such apparently simple phenomena appear amenable to simulation, taking advantage of the continuous increase in hardware power and flexibility/sophistication of simulation environments. We report here the preliminary results of a study on the common features of EEG signals associated to migraine and epilepsy which include: i) a systematic correlation of the spectral content of the EEG signals recorded from individuals with diagnosis of focal and diffused epilepsy and of migraine, and ii) a simulation study of the shift from random to synchronous activity within an artificial Multi Agent System.

CISB Meeting - (Rome, Italy, May 29-30, 2014)

EEG SIGNALS IN EPILEPSY AND MIGRAINE
Analysis and Simulation.

Alessandro Viganò, Neri Accornero and Alfredo Colosimo.

C.I.S.B. - InterDepartmental Research Centre for Models and Information Analysis in Biomedical Systems.

a_mvigano@yahoo.it; neriaccornero@gmail.com; colosimo@caspur.it

Aim of the study

We studied the common features of EEG signals associated to migr aine and epilepsy including: i) correlation of the spectral content of EE G signals recorded from individuals with diagnosis of epilepsy and of migr aine, and ii) simulation of the shift from random to synchronous activity by means of an artificial Multi -Agent System (MAS).

Data Source

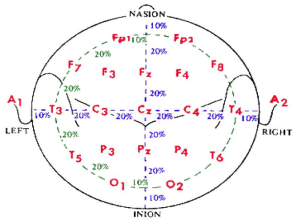


Fig 1. 10-20 EEG montage

The used EEGs were recorded according to 1020 system in the Dept. of Neurological Science of Sapienza – University of Rome, except for the migraine signals coming from the Australian EEG Database (Hunter et al., 2005).

Background

The theory that migraine and epilepsy can be correlated by the abnormal synchronization of neuronal populations was:

1) suggested quite long ago (by H. Jackson). [1]
2) strengthen by populations' studies [2,3]
3) reassessed recently [4].

Statistics about the migraine and epilepsy comorbidity. Fig. 2 shows the prevalence of migraine disease in a European epilepsy-affected population [3]

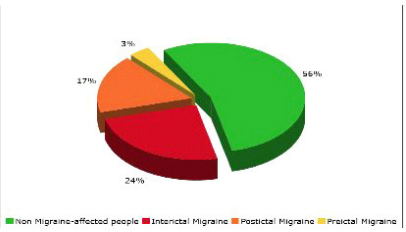


Fig. 2

Cortical Spreading Depression (CSD) was discovered in 1945 [5], in an epileptic model in rabbits. Leao observed this wave moving through the cortex at 3 mm/min speed. The name "depression" is due to the inactivating passage of this wave, since afterward the cortex remains unexcitable. Only in 1994 Lauritzen [6] proposed that CSD was the migraine aura underlying mechanism. In humans CSD was identified in cortex slices taken from epilepsy surgery [7]

Fig 3. shows the CSD spreading mechanism. [8]

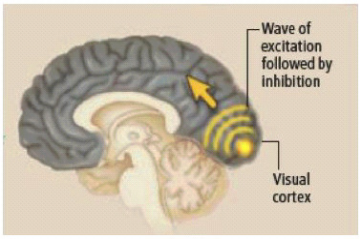


Fig 3.

Data Analysis

Computing Tools: JMP (SAS) statistical package and original macros running on MatLab (Mathwork).

Signals artifacts filtering ⇨ Signals "windowing" in stretches ⇨ power spectra analysis ⇨ PSD data Organization in a 20X50 matrices ⇨ matrices comparison by Pearson's correlation [9]

Pearson's Correlation Coefficent

$$R = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

Simulation Environment

NetLogo

Independent simulations of a synchronizing neural system were carried out by NetLogo. Among Multi-agent system (MAS), NetLogo is a less powerful simulate environment but it appear more flexible than other software. [10, 11].

Results 1: Correlationg EEG signals

		Fp_2	T_4	O_2	T_3	O_1
S1	Fp_1	**0.77**			0.35	0.34
	T_3		**0.64**			0.52
	O_1			**0.77**		
S2	Fp_1	0.79			0.26	0.15
	T_3		0.49			0.22
	O_1			**0.82**		
S3	Fp_1	0.30			-0.04	0.28
	T_3		0.23			0.08
	O_1			0.50		
S4	Fp_1	0.58			0.15	0.40
	T_3		0.22			0.08
	O_1			0.63		
S5	Fp_1	0.53			0.22	0.34
	T_3		0.16			0.14
	O_1			0.64		

Pearson correlations between left and right hemisphere.

S1,S2 = Migraine;
S3 = Diffused Epilepsy;
S4 = Focal Epilepsy;
S5 = Control.

The correlations were calculated from the records of following couples of electrodes: Fp =frontal-pole; T = temporal; O = occipital.

Odd and even suffixes refer to right and left hemispheres, respectively.

Values higher than 0.66 are in bold.

Remarks:

a) all the analyzed signals show a quite synchronous behaviour, between the left and right hemisphere, in fronto-pole and occipital lobes;
b) the signals associated to the migraine diagnosis shows the highest correlation as compared to both the diffused and focal epilepsy cases respectively;
c) the highest synchronous activity is concentrated in the occipita l lobe under all conditions.

Results 2:
Simulating Cortical Spreading Depression

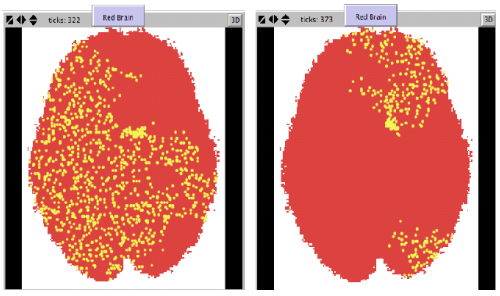


Fig 4. Image from Colosimo, 2008.

Oscillating activity of brain neurons simulated by independent agents. Left and right panels refer to the main peaks of activity within a single oscillation cycle. Each cycle lasted about 10 machine time units, corresponding to about 1 sec, and involved 1500 agents in both hemispheres, whose behaviour was initially random. The oscillating regime apparent in two panels arises after about 350 machine time units. The algorithm used in simulation has been described in [12]).

Conclusions

Our results indicate that by dissecting the time series into a number of subsequent windows the resolution of the method is increased and the presence of time and space ordered activity patterns of neurons from both homo - and contralateral signals can be assessed.

Although still far from conclusive, these results represent an encouraging first step towards the clarification of complex neural pathologies by means of relatively simple and flexible numerical methods.

References

[1] Sacks O. (1992). Migraine. Random House Inc.
[2] Ottman R, Lipton RB. (1996). Is the comorbidity of epilepsy and migraine due to a shared genetic susceptibility? Neurology. Oct; 47(4):918-24
[3] Syvertsen M, Helde G, Stovner LJ, Brodtkorb E. (2007). Headaches add to the burden of epilepsy. J Headache Pain. 8(4):224-30
[4] Rogawski MA. 2008 Common pathophysiologic mechanisms in migraine and epilepsy. Arch Neurol. 65(6):709-14.
[5] Leao AA. (1945). Spreading depression of activity in the cerebral cortex. J. Neurophysiol. 1(7):359-90
[6] Lauritzen M. (1994). Pathophysiology of the migraine aura. The spreading depression theory. Brain. (117):199 -210.
[7] Berger M, Speckmann EJ, Pape HC, Gorji A. (2008). Spreading depression enhances human neocortical excitability in vitro. Cephalalgia. 28(5):558-62.
[8] Dodick DW, Gargus JJ. (2008). Why migraines strike. Sci Am. 299(2):56-63.
[9] Rodgers J, Nicewander W. (1988). Thirteen ways to look at the correlation coefficient. The American Statistician, 42(1):59-66
[10] Brette R, Rudolph M, Carnevale T, Hines M, Beeman D, Bower JM et al. Simulation of networks of spiking neurons: a review of tools and strategies. J Comput Neurosci. 3(3):349-98.
[11] Wilensky U. (2009). NetLogo – Home page. http:// ccl.northwestern.edu/netlogo
[12] Colosimo A. (2008) Biological simulations by autonomous agents: two examples using netLogo enviroment. Biophysics and Bioengineering Letters. 1(3):40 -50.

7th International Conference on Acoustic Emission 12-15 Sept 2012 Granada, Spain

RQA AND EARTHQUAKE PRECURSORS

Giovanna Zimatore, Maurizio Poscolieri, Claudio Rafanelli

ABSTRACT

Non linear time series techniques are largely being applied in different disciplines, such as Earth Sciences, Astrophysics, Engineering, Economy, Physiology, and Neurosciences. One, in particular, Recurrence Quantification Analysis (RQA), appears quite promising for the analysis of complex systems. Moreover, it has recently been applied to investigate both acoustic emissions from rocky specimens and the dynamics of complex seismic processes.

Whenever two objects are put in contact and let them slide with respect to each other, friction induced vibrations may occur. Examples are represented by active faults within seismogenic zones, narrow- banded noise of train wheels running along tight curves, friction in bearings, and micro-scale events in molecular physics.

In this framework, our work describes the application of RQA to the Passive Acoustic Emission (AE) signal released, at ultrasonic frequencies, by stressed rocks in the Earth's crust above a given threshold (event). The data record is represented by AE time series gathered, with 30 sec. of sampling rate, at the Valsinni (Potenza, Italy) Orch (Perugia, Italy) and Peteroa Volcano (Argentinean Andes) monitoring stations. In these sites AE data were collected by piezoelectric transducers, working at two ultrasonic frequencies, stuck to a rock. This way, a huge amount of data is available but the AE signal amplitude varies with the acoustic impedance, related to local rocks stress conditions and particularly sensitive to fracture density and water content. Our aim is to identify few descriptors that can explain the main characteristics of the AE signals and identify anomalies to be related to crustal stress modifications or, as in the Peteroa case study, paroxysmal volcanic activities or Earth's tides imprints.

RQA AND EARTHQUAKE PRECURSORS

Giovanna Zimatore, Maurizio Poscolieri, Claudio Rafanelli
CNR-IDASC – Institute of Acoustics and Sensor "Orso Mario Corbino" – ICES Group
Area della Ricerca "Roma - Tor Vergata", via del Fosso del Cavaliere, 100, Rome, Italy

RQA is applied in this work to analyze time series of acoustic emission (AE) triggered by Earth's crust phenomena, so to detect feasible precursors of catastrophic events.

The AE records of rms values were acquired at different sites considering two ultrasound frequencies (LF= 25kHz e HF=150kHz) with 30-sec sampling rate.

Sites:Valsinni (Matera, Italy), Orchi (Perugia, Italy) and a location close to the Peteroa volcano (Argentina),

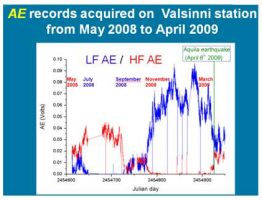
The preliminary results of the application of RQA to the AE data are described and appear suited to stress relationships with an impending catastrophic event.

RQA is applied in this work to analyze time series of acoustic emission (AE) triggered by Earth's crust phenomena, so to detect feasible precursors of catastrophic events. Quantitative changes in dynamics of temporal distribution, loss of synchronization of dynamic mechanism or spatial irregularities occurring in time can be appreciated.

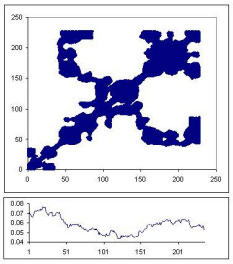
Acoustic Recording probe (Orchi,Italy)

Operative AE stations

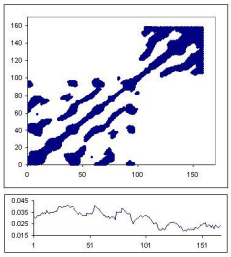
AE records acquired on Valsinni station from May 2008 to April 2009



Recurrence plot of AE recorded in Valsinni station (Italy)



before L'Aquila earthquake (9th 2009)
%Rec=25.8, %Det=90.4



after L'Aquila earthquake
%Rec=16.1, %Det=71.61

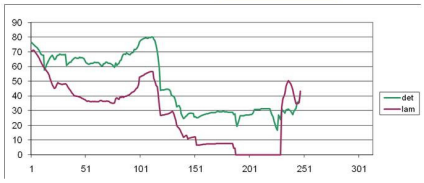
By RQC from http://homepages.luc.edu/~cwebber/

Recurrences analysis in sliding windows of AE recorded in Orchi station (Italy)

Start	End	Transition	Sismic Event
Epoch 1 06/02/2013	Epoch 251 23/02/13	Epoch 115 13/02/2013	Site: Sora (M=4.8) Date: 16 /02/2013

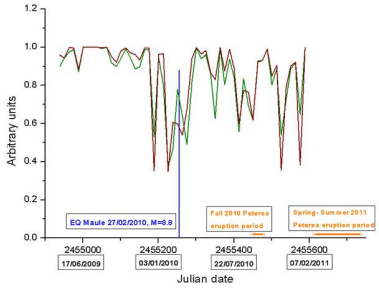
Epoch :150 points
Shift: 1 point

Distance Sora-Orchi:162 km



By RQE from http://homepages.luc.edu/~cwebber/

Recurrences analysis in sliding windows of AE recorded in Peteroa station (Argentina)



RQA parameters calculated from the HF AE series: the green line points out the *%DET* descriptor, while the brown line the *%LAM* (51 non-overlapping epochs; 1 epoch=300 hours). From JAE_2012

Conclusions

- These results suggest a relation between Det decreasing and energy release during an earthquake.
- In our preliminary studies, a correlation between fractal dimension and RQA parameters (%Rec and %Det) was observed.
- In the future, we will apply PCA on RQA parameters to search possible correlations with events occurring in regions included in a radius of 100, 150, 200, 250 km from AE stations.
- New AE stations and new probes will be closely realized in the south of Italy and additional data will be available

References

- G. Paparo and G.P. Gregori, 'Multifrequency acoustic emissions (AE) for monitoring the time evolution of microprocesses within solids, reviews of quantitative nondestructive evaluation',22, AIP Conference Proceedings, D.O. Thompson, and D.E. Chimenti (eds.), (2003), pp. 1423-1430.
- C.L. Webber and J.P Zbilut., 'Dynamical assessment of physiological systems and states using Recurrence Plots Strategy', J. Appl. Physiol., 76(2) (1994), 965-973.
- N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan and J. Kurths, 'Recurrence plot based measures of complexity and its application to heart rate variability data', Phys. Rev. E, 66(2) (2002), 026702.1-026702.8.
- A.N. Gorban, E.V. Smirnova and T.A. Tyukina, 'Correlations, risk and crisis: from physiology to finance', Physica A, 389 (2010), 3193-3217.
- T. Chelidze and T. Matcharashvili, 'Complexity of seismic process; measuring and applications - A review', Tectonophysics, 431 (2007), 49-60.
- M. Takemoto, H. Nishino and K. Ono, 'Wavelet Transform - Applications to AE Signal Analysis', in: Acoustic Emission - Beyond the Millennium, Elsevier, (2000), pp. 35-56.
- Maurizio Poscolieri, Claudio Rafanelli and Giovanna Zimatore Viable Precursors of Paroxysmal Phenomena as Detected by Applying RQA to Acoustic Emission Time Series. Journal of Acoustic Emission, Volume 30 (2012) 29-39
- Giovanna Zimatore, Domenico Stanzial and Maria Patrizia Orlando "Otoacoustic Emissions" Chapter 9 of Acoustic Emission - Research and Applications, Dr. Wojciech Sikorski (Ed.), ISBN: 978-953-51-1015-6, (2013)

Index of Authors

Many sincere THANKS to

the Speakers,
for their contribution to the Proceedings volume;

the former CISB technical staff,
for their help in organizing the Meeting.