
This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

[http://eprints.gla.ac.uk/143594/](http://eprints.gla.ac.uk/143594/)

Deposited on: 27 July 2017

Enlighten – Research publications by members of the University of Glasgow
[http://eprints.gla.ac.uk](http://eprints.gla.ac.uk)
Analysis of the fluctuations of the tumour/host interface

Edoardo Milotti\textsuperscript{a}, Vladislav Vyshemirsky\textsuperscript{b}, Sabrina Stella\textsuperscript{a,\ast}, Federico Dogo\textsuperscript{a}, Roberto Chignola\textsuperscript{c}

\textsuperscript{a}Department of Physics, University of Trieste, Via Valerio 2, I-34127 Trieste, Italy.
\textsuperscript{b}School of Mathematics and Statistics, University of Glasgow, G12 8QQ Glasgow, Scotland, UK
\textsuperscript{c}Department of Biotechnology, University of Verona, Strada Le Grazie 15 - CV1, I-37134 Verona, Italy

Abstract

In a recent analysis of metabolic scaling in solid tumours we found a scaling law that interpolates between the power laws $\mu \propto V$ and $\mu \propto V^{2/3}$, where $\mu$ is the metabolic rate expressed as the glucose absorption rate and $V$ is the tumour volume. The scaling law fits quite well both \textit{in vitro} and \textit{in vivo} data, however we also observed marked fluctuations that are associated with the specific biological properties of individual tumours. Here we analyse these fluctuations, in an attempt to find the population-wide distribution of an important parameter ($A$) which expresses the total extent of the interface between the solid tumour and the non-cancerous environment. Heuristic considerations suggest that the values of the $A$ parameter follow a lognormal distribution, and, allowing for the large uncertainties of the experimental data, our statistical analysis confirms this.

\textit{Keywords:} fluctuations in biological systems, scaling laws, lognormal distribution, cancer metabolism, cancer growth

\textsuperscript{\ast}Corresponding author

\textit{Email address:} sabrina.stella@ts.infn.it (Sabrina Stella)

\textit{Preprint submitted to Physica A June 13, 2017}
1. Introduction

The allometric laws of biology are power laws that describe in simple and universal terms some of the features that are common to the vast diversity of living organisms [1, 2, 3]. Their simplicity hides the ambitious aim of encapsulating in a single equation the huge variability and diverse dynamics that belong to living systems. Sometimes, their purely empirical character has produced controversy, and some of these laws are not universally accepted [4, 5, 6]. Kleiber’s law is one such scaling law: it states that the basal metabolic rate of an organism of mass \( M \) scales as \( M^{3/4} = M^{0.75} \) instead of the naive estimate \( M^{2/3} \approx M^{0.67} \), and it has both supporters and opponents. Kleber’s law has long been mysterious and only in recent years it has found the backing of a biological argument based on fractal structure of microcirculation [8], which is however still controversial [9, 10, 11]. Nonetheless, the law has been applied also outside the domain of validity of the supporting argument, as in the case of the metabolism of solid tumours [12]. Indeed, microcirculation in tumour tissues is quite different from normal tissues, with an ensuing wide-ranging variability of cellular phenotypes and local microenvironments. However, it is well-known that there is a strong correlation between tumour microenvironment and aggressiveness, and this provides a strong motivation to study the microscopic origin of any metabolic scaling law that may arise in this context.

We have recently proposed a metabolic scaling law for solid tumours that describes quite well both in vitro and in vivo experimental data (see figure 1) [13]. Its peculiarity is that it has been obtained with the help of a computational description of metabolism at the single-cell level which provides the link with the microscopic features of tumours [14, 15, 16]. The law is not quite a power law, and it is given by the following expression:

\[
\mu = \left( \frac{c}{v_c} \right) \frac{3\lambda V}{3\lambda + (V/A)^{1/3}}
\]

where \( \mu \) is the metabolic rate expressed as the glucose absorption rate, \( V \) is the tumour volume, \( v_c \) is the mean cell volume, \( c \) is the cell-specific consumption rate
Figure 1: Metabolic rate data $\hat{\mu}_N = \mu/\eta A^{1/3}$ vs. $V/A(\mu m^3)$. The figure shows a breakdown of the data originally presented in [13]. Upper panel: tumor spheroids cultured in vitro using different cell lines. Lower panel: data from human tumours, that include breast, uterine and ovarian carcinomas, melanomas, thyroid carcinomas, colon and lung carcinomas. All of these human tumours correspond presumably to different cell-specific metabolic rates $\eta$. The dotted line is $\hat{\mu}_N = V$, while the dashed line is $\hat{\mu}_N = V^{2/3}$, which are the extreme behaviours derived from equation (1).

of a single cell [13], and A is a non-dimensional parameter whose meaning will be clarified below. The $\lambda$ parameter is a characteristic length in an exponential law that describes the decay of the fraction of live cells with increasing distance from the blood vessels (see [10] for details). The ratio $\eta = c/v_c$ in equation (1) is independent of cell size, since $c$ is roughly proportional to cell volume. Equation (1) shows that the metabolic rate interpolates between the two power-laws

$$\mu \approx \left( \frac{c}{v_c} \right) V \quad (2)$$
at small volume and
\[ \mu \approx \left( \frac{c}{v_c} \right) \frac{3\lambda}{(3/4\pi)^{1/3}} V^{2/3} \]  
(3)

at large volume, and if we take the middle value of the exponent (0.84) we note that it is not very far from the exponent of Kleiber’s law.

Turning to parameter A, we note that is related to the total extent of the interface between solid tumour and surrounding environment, see fig.2 if \( x \) is a diameter (chosen between two recognizable features of the tumour) then the total interface area \( S \) scales as \( S = Ax^2 \) (see also ref [13]). The analysis of existing data – both \textit{in vitro} and \textit{in vivo} – points to a strong dependence of the metabolic rate on \( A \). In particular, we attribute the spread of data about the theoretical expression shown in figure 1 to the fluctuations of \( A \) in populations of histologically similar tumours. Since the interface area of a tumour with the normal tissue environment influences both its growth rate and its metastatic potential, these fluctuations are clinically relevant.

In this paper we argue that the probability distribution of \( A \) is lognormal, and we show that the experimental data support this conjecture.

2. Theory

The \( A \) parameter is directly related to the complexity of the interface between tumour and environment (see figure 2), and the interface growth can be pictured as a gradual buildup of new features that appear on already existing elements, in a way that is multiplicative rather than additive – as in a Kolmogorov process [17]. Such a process has already been assumed in simple simulations of biological growth based on analogs of diffusion limited aggregation [18, 19], and if growth can be mapped on a Kolmogorov process this leads to a lognormal distribution [17, 20, 21]

\[ p(A; m, \sigma) = \frac{1}{A\sqrt{2\pi}\sigma^2} \exp \left[ -\frac{(\ln A - m)^2}{2\sigma^2} \right] \]  
(4)

where \( m \) and \( \sigma \) are the parameters that define the shape of the distribution. This type of growth has already been used in an attempt to explain the complex
structure of human lungs [22]. In this section we demonstrate the plausibility of the lognormal conjecture and show how to obtain values of \( A \) from the data shown in figure [1].

Figure 2: Schematic view of the tumour-environment interface in solid tumours and in tumour spheroids. 1. Cross section of a solid tumour at low magnification. Live tumour cells (T) wrap around blood vessels (BV) to form tumour cords. Because of the limited diffusion of nutrients, tumour cells that are distant from blood vessels die and become necrotic (N). 2. Enlarged view of a single tumour cord. The density of live cells decreases at increasing distance from blood vessels as dead cells mix with live cells. 3. A multicellular tumour spheroid (MTS) – an important in vitro model of avascular tumours – surrounded by culture medium (M) at low magnification. Live tumour cells (T) proliferate in the external layers, while dead cells form a necrotic core (N). 4. Enlarged view of the MTS section shown in 3: live cells mix with dead cells just as in real tumours. In both real tumours and MTS live cells wrap and fold around the nutrient supply system, i.e. blood vessels and the external environment, respectively. The value of \( A \) is determined by the interface between the bulk of the living tumour cells and the non cancerous environment, and this includes the network of blood vessels that supply the tumour with nutrients. Therefore \( A \) is expected to be much higher in vascularised solid tumours than in avascular tumour spheroids.
2.1. The distribution of the $A$ parameter

Figure 2 contains two classes of solid tumours: solid tumours \textit{in vivo} and tumour spheroids \textit{in vitro}, and we consider the lognormal arguments separately for each class.

2.1.1. The case of \textit{in vivo} solid tumours

In this case the conjecture that the $A$ parameter follows a lognormal distribution is made plausible by the following considerations. In a solid tumour with total volume $V$ and total interface area $S$, the interface corresponds to all the surface where tumour cells exchange oxygen, nutrients and metabolites with the environment, and can be identified with the set of blood vessels that traverse the tumour mass. This corresponds to a large area, and if we make the simplifying assumption that the radius $r$ of blood vessels is fixed, then it is given by $S \approx 2\pi r L$, where $L$ is the total length of the blood vessel network.

When the linear size of the volume increases by a factor $\alpha$, we find that the volume increases by a factor $\alpha^3$ and the surface area by a factor $\alpha^2$, however this also means that the total length of the blood vessel network also increases by the same factor $\alpha^2$. This tells us that blood vessels sprout new vessels, and the contribution of these new blood vessels to the total network length is the difference between the quadratic growth and a simpler linear growth

$$\alpha^2 L - \alpha L = \alpha L (\alpha - 1)$$

The same reasoning works also in the more general case where the growth of the interface area is not exactly quadratic, but has a fractional exponent $S' = \alpha^\beta S$, then the contribution of new vessels to the growth of the network length is

$$\alpha^\beta L - \alpha L = \alpha L (\alpha^{\beta-1} - 1)$$

Whatever the case, there cannot be a continuous growth, and the evolution of the tumour mass must be punctuated by discrete sprouting events, with a random distribution of $\alpha$ at each step. We can also write

$$S = \frac{dV}{dx} = \frac{d(\alpha^3 V_0)}{d(\alpha x_0)} = 3\alpha^2 \frac{V_0}{x_0} = 3\alpha^2 A_0 x_0^2$$

6
in the case of growth over many sprouting events, where subscripted values are fixed at a given time $t_0$, or also

$$ S = \frac{dV_{n+1}}{dx} = \frac{d(\alpha_{n+1}^3 V_n)}{d(\alpha_{n+1} x_n)} = 3\alpha_{n+1}^2 \frac{V_n}{x_n} = 3\alpha_{n+1}^2 A_n x_n^2 $$

(8)

for growth between two sprouting events. Then we see from this equation that the actual value of $A$ at any given time fluctuates about an average, and that it is given by

$$ A = \alpha^2 A_0 = \prod_k \alpha_k^2 A_0 $$

(9)

where $\alpha$ is the combined, multiplicative result of many blood vessel sprouting events. This means that

$$ \ln A = 2 \sum_k \ln \alpha_k + \ln A_0 $$

(10)

so that $\ln A$ is a (large) sum of variates, and if the assumptions of the central limit theorem hold, then $A$ has a lognormal distribution.

2.1.2. The case of in vitro tumour spheroids

The previous considerations do not apply to avascular tumours like the tumour spheroids, however we can pinpoint processes that contribute to a similar multiplicative structure in the determination of the actual interface area. Indeed, it has long been known that tumour spheroids grown in identical conditions can have widely differing sizes at saturation [23, 24], and that this is closely related with the spheroid structure, in particular with the formation of the necrotic core: this fact underscores the importance both of the random events at the single-cell level – the only difference in the development of these spheroids – and of the structural elements. Another hint is provided by observations of the surface of tumour spheroids, which is often rough and marked by hills and valleys; this fact has recently been exploited in a study of the delivery of nanoparticles into tumour tissue [25]. Irregularities such as these are again the result of discrete and at least partially random events, like cell death in the necrotic core and the mitoses in the outer cell layer of tumour spheroids, and
the previous reasoning applies again – although we expect to find a lognormal
distribution with different parameters.

2.2. Finding $A$ from metabolic data

Here we remark that we can use the metabolic scaling law to determine $A$
when the other parameters are fixed

$$A \approx \frac{V}{(3\lambda)^3(\eta V/\mu - 1)^3}$$

and we do so both for cultured tumour spheroids, where it is easier to estimate
$\eta$, and for human tumours, using an average $\eta$. For a given histological type,
both $\eta$ and $\lambda$ are nearly fixed, and here we take the values of $\lambda$ estimated in
[13] for tumour spheroids and for human tumours. Because of the difficulty of
obtaining valid data, we also use the same datasets of [13], that we selected
after an extensive search of the existing literature. And indeed, while there are
many data available, their usability is limited by the following requirements:

- we utilise glucose absorption to define the metabolic rate; therefore all the
data used in equation (1) or (11) must refer to glucose;

- it is difficult to find combined values of the parameters for the same cell
line;

- we cannot use data expressed in arbitrary units. This rules out many data
like those obtained with uncalibrated imaging techniques, those obtained
with radioactive markers without any indication of the specific activity
of labelled compounds, and data normalised per gram of tumour tissue
without any reference to the total amount of tissue used.

3. Calculation

We used the procedure of section 2.2 to obtain the $A$ values corresponding to
the data shown in figure [1]. Since these data are affected by rather large uncer-
tainties, the complete probability density function (pdf) of each measurements
is the convolution of the lognormal distribution with a normal distribution (we make the usual assumption that measurement uncertainties have a normal distribution):

\[
p(A, \sigma_0; m, \sigma) = \\
\int_{-\infty}^{A} \frac{1}{(A - x)\sqrt{2 \pi \sigma^2}} \exp \left\{ -\frac{[\ln(A - x) - m]^2}{2 \sigma^2} \right\} \frac{1}{\sqrt{2 \pi \sigma_0^2}} \exp \left( -\frac{x^2}{2 \sigma_0^2} \right) \, dx \quad (12)
\]

where \( \sigma_0 \) is the measurement error. Then the complete log-likelihood is

\[
\ln L(D; m, \sigma) = \sum_i \ln p(A(i), \sigma_0(i); m, \sigma) \quad (13)
\]

where \( D = \{A(i), \sigma_0(i)\}_{i=1,n} \) is the set of all \( n \) data values and their measurement errors. The evaluation of the log-likelihood requires the numerical evaluation of all the individual likelihoods, since expression (12) does not have a closed analytical form. The log likelihood (13) is only approximate because the large measurement errors on \( A \) cannot be Gaussian – \( A \) is a non-negative parameter – and moreover it depends strongly on the specific measurement errors \( \sigma_0(i) \). Moreover, we have only incomplete information on the measurement themselves and could not obtain good error estimates for many datapoints. All of this leads to large numerical uncertainties and to inconclusive results when using the log-likelihood (13), so that eventually we opted to fit the values \( A \) deduced with (11) with the lognormal pdf (4), and our results are also shown in figures 3 and 4.

We tested the validity of the lognormal hypothesis by comparing it with a highly flexible phenomenological model, a Gaussian mixture model with a variable number of components. We did this in a Bayesian framework, and we ranked the different hypotheses using marginal likelihoods, both for cultured spheroids and for solid tumours.

In the case of cultured spheroids we considered a total of 35 alternative models, the lognormal model and 34 Gaussian mixture models with a number of components ranging from 1 to 34. We defined only 34 mixture models because the data set contains 34 samples, and therefore any more complex model would
Figure 3: Pdf of the values of $A$ in the case of the aggregated tumour spheroid data (spheroids from the 9L, MR1 and Rat-T1 cell lines): the original data points are shown as a “rug plot” (the positions marked by the vertical bars in the lower part of the plot). In addition to the glucose uptake values given in [13] we have used $\eta = 6.5 \times 10^{-12}$ and $\lambda = 93.9 \ \mu m$ for 9L spheroids, $\eta = 4.2 \times 10^{-12}$ and $\lambda = 123.6 \ \mu m$ for Rat-T1 spheroids; $\eta = 8.9 \times 10^{-12}$ and $\lambda = 91.0 \ \mu m$ for MR1 spheroids. The solid curve is the lognormal fit, the dashed line is the empirical smoothed kernel density, and the dotted line is a power-law with exponent -1.3.

Figure 4: Pdf of the values of $A$ in the case of human solid tumours: the original data points are shown as a “rug plot” (the positions marked by the vertical bars in the lower part of the plot). In addition to the glucose uptake values given in [13] we have used $\eta = 3.7 \times 10^{-12}$ and $\lambda = 99.4 \ \mu m$. The solid curve is the lognormal fit, the dashed line is the empirical smoothed kernel density, and the dotted line is a power-law with exponent -1.3.

certainly be less plausible than the one with 34 components. Thus, we have the following likelihoods for individual data samples:
Lognormal model.

\[ p(A|m, \sigma) = \logN(A; m, \sigma) \]  \hspace{1cm} (14)

Model 1.

\[ p(A|\theta) = \sum_{i=1}^{1} \alpha_i N(A; m_i, \sigma_i) \quad \theta = (\alpha_1, m_1, \sigma_1) \]  \hspace{1cm} (15)

\ldots

Model 34.

\[ p(A|\theta) = \sum_{i=1}^{34} \alpha_i N(A; m_i, \sigma_i) \quad \theta = (\{\alpha_i\}_{i=1,34}, \{m_i\}_{i=1,34}, \{\sigma_i\}_{i=1,34}) \]  \hspace{1cm} (16)

where \( \logN(A; m, \sigma) \) is the lognormal pdf with parameters \( m \) and \( \sigma \) as above, and \( N(A; m, \sigma) \) is the normal pdf with parameters \( m \) and \( \sigma \).

We introduced weakly informative priors for these models. For the lognormal one, \( m \) and \( \sigma \) were assigned uniform priors from 0 to 200. For the mixture models of \( M \) components, we assigned \( M - 1 \) uniform priors for \( \alpha_i \) in the interval 0 to 1, added an extra improper prior for \( \sum_{i=1}^{M-1} \alpha_i < 1 \), define \( \alpha_M = 1 - \sum_{i=1}^{M-1} \alpha_i \), assigned all \( m_i \) and \( \sigma_i \) to be independent and uniformly distributed from 0 to 200, and finally we added an improper prior for \( m_1 \leq m_2 \leq \cdots \leq m_M \).

To rank alternative hypotheses, we need to estimate marginal likelihoods for the alternative models. We estimated these likelihoods using the Thermo-dynamic Integration method proposed by [26], and obtained the results shown in figure 5. Together with the Bayes factor for preferring the lognormal model over any of the mixture models having the value of \( 4.45 \times 10^{18} \), this indicates decisive evidence that the lognormal model should be preferred over any mixture of Gaussians. Our solid tumour dataset has just as many samples as the tumour spheroid dataset and we repeated the same kind of analysis; we used uniform priors from 0 to 200000 for the \( m \)’s and \( \sigma \)’s, since we expect a wider range of \( A \). In this case the lognormal hypothesis ranks only second (see figure 6), however we note that the solid tumour data are much more dependent on the assumptions of metabolic consumption rate, etc., and therefore they carry
With them potentially large systematic errors. Moreover, the Gaussian mixture model is purely phenomenological and does not rely on laws of scaling in Biology. The better performance of this mixture model can be explained by its relative simplicity while preserving good flexibility for matching experimental data. Notice that the lognormal model fares better than the large majority of the highly flexible Gaussian mixture models and in contrast to all of them it is biologically motivated.

Figure 5: Ranking of hypotheses in the case of in vitro tumour spheroids. The numbers on the x axis mean the number of mixture components, the y axis is the value of the marginal likelihood.

4. Results and Discussion

As noted above, the A parameter is related to the overall shape of the tumour. It is one of the factors that set the timescale of tumour growth and has an importance of its own, as it determines both the total tumour volume and the size of the interface between tumour and nutrient-supplying blood vessels. This interface also regulates the influx of drugs and influences the overall metastatic potential because it is the place where tumour cells can enter the blood stream. The heuristic considerations discussed above suggest that tumour spheroids and
solid tumours in vivo both have lognormal distributions, although with different parameters. The actual observations indicate that the distinction is correct, and show that the $A$ parameter becomes quite large in vascularised tumours in vivo because of the fractal nature of the capillary network that feeds tumours [27]. The statistical analysis lends a support to the lognormal conjecture that is strong in the case of in vitro tumors, but it is still inconclusive for in vivo tumours and a compelling statement requires additional data.

Interestingly, when the $\sigma$ parameter is smaller than $m$ the tail of the lognormal distribution provides a reasonable approximation of a power-law distribution over several orders of magnitude, a fact that has prompted [20, 21] to argue that the ubiquitous power-law distributions may actually stem from underlying lognormals. If this turned out to be true also in this case, then in a certain subrange of values the $A$ parameter would display a sort of scale invariance over the population of solid tumours – this is also illustrated in figures 3 and 4 where we plotted a power law $\sim 1/A^{1.3} \approx 1/A^{4/3}$ to show the kind of scale-invariance
that we would obtain with the present data analysis.

4.1. Conclusions

The $A$ parameter sets the extent of the interface area, and for this reason it has an obvious correlation with tumour aggressiveness and with the structure of the tumour microenvironment, and a detailed knowledge of its distribution can yield important statistical estimates on the overall behaviour of a population of solid tumours. This requires a protocol for measuring $A$ in a large number of cases, and we expect that studies of $A$ for different histological tumour types can lead to different parameter sets. Finally, the global statistical parameters could be used to assess the eventual prognosis with a given dynamics of $A$ in an individual tumour.

Data Availability

The data used in this work are taken from our previous work, and they are listed in [13], which is available as an open access article:

http://www.nature.com/articles/srep01938.

Competing Interests

The authors declare no competing interests.

Funding

This work has been supported by a grant FRA 2013 from the University of Trieste.

References


tal tumors from as few as two determinations of tumor size: implications
for clinical oncology, IEEE transactions on biomedical engineering 52 (5)

Tumour-on-a-chip provides an optical window into nanoparticle tissue

[26] N. Friel, A. N. Pettitt, Marginal likelihood estimation via power posteriors,
Journal of the Royal Statistical Society: Series B (Statistical Methodology)

of tumor vascular architecture in nutrient and drug delivery: an invasion