

The Cancer Cell Spread and Metastasis: A Cell and Life Perspective

Haowei Ti^{1*}, Yu Ti², Xijie Zhao³

Abstract

Background: During the clinical process, it is difficult to judge the life cycle of cancer patients, but based on the spread speed of cancer cells, the approximate cycle can be calculated.

Purpose: This study uses conventional mathematical statistics to track the spread of cancer cells in a certain number of cancer patients and uses mathematical models to calculate the approximate life cycle of the patients.

Method: This article randomly selects the cancer cell proliferation of liver cancer as the main target from the overall population and then uses the impact of the cancer cell proliferation rate on the life cycle of 20 gastric cancer patients to obtain 50 data samples. The sample capacity is 20. This article estimates the population mean parameter based on the sample mean. It is based on describing the sample data and making inferences in the form of probabilities about the unknown quantitative characteristics of the statistical population. Rather, the sample data are inferred from observations of a random process over a limited period.

Result: This article uses samples to infer (estimate) the population distribution parameters in many ways. This article will assume that the overall distribution obeys a certain known probability distribution, but some parameters of the distribution are uncertain.

Conclusion: The spread speed of cancer cells does not completely affect the patient's life cycle. If resistance to targeted drugs can be resolved, cancer cells and normal cells can coexist, and patients can continue to survive.

Keywords: Cancer cells; Life cycle; Mathematical model; Lung cancer.

Introduction

In the clinical process, the life cycle of cancer patients is a very taboo matter. We often encounter cancer patients passing away, but it is difficult to know their life cycle. We may make a rough judgment based on experience, but the more specific details are unknown. We use mathematical methods to calculate and evaluate each tumor size change and the physical changes caused by the spread of cancer cells. The research on cancer cell spread involved in this article is aimed at patients with stage II and stage III cancer. The research in this article is limited to the normal spread of cancer cells and does not involve the intervention of targeted drugs and radiotherapy/chemotherapy.

Primary tumors metastasize to distant organs in three main ways:

1. By passing through the circulatory (blood) system.
2. Through the lymphatic system.

Affiliation:

¹Doctor of Clinical Medicine, Rey Juan Carlos University, Madrid, 28028, Spain

²Qiaotou Middle School, Dongguan, 523000, China

³Dalang Middle School, Dongguan, 523000, China

*Corresponding author:

Haowei Ti, Doctor of Clinical Medicine, Rey Juan Carlos University, Madrid, 28028, Spain

Citation: Haowei Ti, Yu Ti, Xijie Zhao. The Cancer Cell Spread and Metastasis: A Cell and Life Perspective. *Journal of Cancer Science and Clinical Therapeutics*. 10 (2026): 28-34.

Received: February 09, 2026

Accepted: February 17, 2026

Published: March 31, 2026

3. Enter the abdominal cavity and thoracic cavity through the body wall (implantation metastasis).

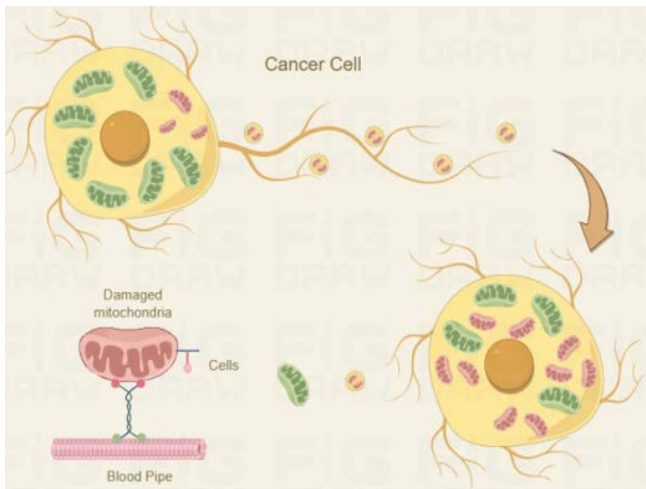


Figure 1: Cancer cell depart

The purpose of this article is to examine the changes in cancer cell spread and life cycle in stage II and III patients. The circulatory system is the primary route of spread to distant organs, while lymphatic vessels provide access to regional lymph nodes. After this, the cancer cells usually spread through the blood. Although body fluid transmission is the most common route, hematogenous spread via lymph or blood depends on the origin and location of the primary tumor. Lung cancer cells spread through lymph and blood. If cancer cells spread throughout the body, they must jump over or bypass nearby normal cells to accelerate their spread. They attach to other cells and the extracellular matrix by remodeling the cytoskeleton, and through proteins on the surface of the cytoplasmic membrane. Cancer cells can spread forward by extending the bulge of the cell forward and allowing the rear to follow. The cancer cell crawls forward until it hits an obstacle it cannot bypass. The enzymes secreted by cancer cells contain a group of enzymes called matrix metalloproteases (MMPs). These enzymes act as "scissors" to snip out proteins that block the movement of metastatic cancer cells. Once across the basement membrane, cancer cells can metastasize throughout the body in a variety of ways. For example, they can squeeze through the spaces between vascular endothelial cells and enter the bloodstream or body fluids for diffusion or transfer. Once in the bloodstream, cancer cells travel through the circulation until they find a suitable place to take up residence, and then they re-enter the tissue. They begin to grow and multiply in new locations, forming new tumors.

Cancer cell metastasis formation is a very inefficient process but is responsible for most cancer-related deaths. This is because millions of cells can leave tumors every day.

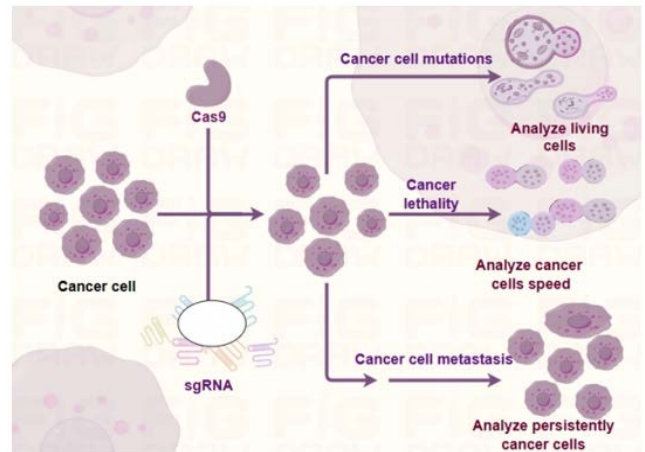


Figure 2: Cancer cell metastasis formation

Differences in the Spread Speed of Lung Cancer

The growth of tumors in the body follows linear growth after exponential growth; the flow rate of cancer cells in human body fluids also follows linear growth.

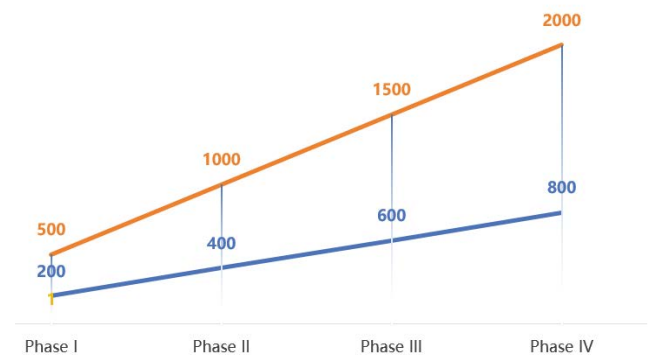


Figure 3: Cancer growth curve

Based on the characteristics of the tumor growth curve, the Gompertz function with similar shape characteristics was used for fitting and description.

$$\frac{dE}{dt} = E * (\alpha - \beta * \ln(\frac{E}{E_0})) \tag{1}$$

The threshold value of tumor mass W_{th} , at which the growth of the tumor changes from an exponential process to a linear process (i.e., from a 1st-order kinetic process to a 0th-order kinetic process), can be described by the following differential equation:

$$\begin{aligned} \frac{dw(t)}{dt} &= \lambda_0 * w(t), w(t) \leq W_{th} \\ \frac{dw(t)}{dt} &= \lambda_1, w(t) > W_{th} \\ w(0) &= w_0 \\ \lambda_1 * W_{th} &= \lambda_1 \end{aligned} \tag{2}$$

Among them,

$w(0)$ represents the mass of the tumor at the time of inoculation.

λ_0 represents the growth rate constant in the exponential growth phase, (1st-order kinetic rate constant);

λ_1 represents the growth rate constant in the linear growth phase, (0th-order kinetic rate constant);

W_{th} represents the threshold value of tumor mass.

For computational reasons, it is more convenient to describe using only a single differential equation, especially for the metastatic speed of cancer cells to be introduced, so the following equation is used as an approximation of the above equation:

$$\frac{dw(t)}{dt} = \frac{\lambda_0 \cdot w(t)}{\left(1 + \left(\frac{\lambda_0}{\lambda_1} \cdot w(t)\right)^\Psi\right)^{\frac{1}{\Psi}}} \quad (3)$$

$w(0)=w_0$

When Ψ is large enough, the above equation is a good approximation of the original equation.

When the tumor mass $w(t)$ is less than W_{th} , the term $\{\lambda_0/\lambda_1 w(t)\}^\Psi$ becomes very small, and relative to the 1 in the denominator, this term can be ignored, so the cancer cell growth rate is approximately $\lambda_0 w(t)$.

When the tumor mass $w(t)$ is greater than W_{th} , the term $\{\lambda_0/\lambda_1 w(t)\}^\Psi$ becomes very large, and relative to the 1 in the denominator, the 1 in the denominator can be ignored, so the cancer cell growth rate is approximately λ_1 .

Lung cancer is divided into squamous cell carcinoma, adenocarcinoma, large cell lung cancer, and small cell lung cancer, and the first three are combined into non-small cell lung cancer. However, many lung cancers contain several different cancer types at the same time, and the proportions are different, so the classification is different. For example, if a tumor contains both squamous cell carcinoma and adenocarcinoma, the proportion of one component less than 10% will be ignored. If both exceed 10%, it is separately classified as adenosquamous carcinoma. And this 10% is artificially defined.

Secondly, even if it is a single cancer type, the performance of different patients still varies greatly. Some people have very large tumors and have almost no symptoms at all, while others have very small tumors but have severe symptoms. Some people's tumors mainly invade locally, while some primary tumors are small but have metastasized everywhere. The third is the speed of spread. It involves many levels. It can refer to the speed at which the tumor doubles in size. It can also refer to a tumor's ability to readily metastasize to distant sites. Cancer type is one of the variables that determines these abilities, but it is not the only, and probably not the most important, variable.

Now suppose we have some observations of the variable X . The set of these observations is represented by the symbol $D=\{x(1), x(2), \dots, X(N)\}$. These observed values are all obtained from the same probability distribution $P = (X; \theta)$, and these samples are obtained independently, that is, each sample value does not depend on other sample values. We can say that these samples are independent and identically distributed. The probability distribution of each variable in any set of random variables is the same, and these random variables are independent of each other.

Sample Set

This article mainly studies the impact of cancer cell expansion speed on the life cycle of cancer patients. Then the data on the proliferation speed of all cancer cells is an overall body. The data distribution to which all cancer cells spread data belongs is called the overall distribution, in which each person's cancer cell spread data, that is, a single data, is called an individual. Selection sampling means randomly selecting some individuals from the population and then obtaining the data of these sampled individuals. The result of one sampling is called a sample. We collected 20 pathologically confirmed pulmonary LCNEC patients who were diagnosed in the hospital from January 2015 to June 2020, including 15 males and 5 females, aged 49-78 years old, with an average age of 62.82 years. There were 15 smoking patients, with a smoking index of 0-3,000, and an average smoking index of 931.82.

The staging standard adopts the 2009 7th edition of the tumor-node-metastasis (TNM) staging system of the International Association for the Study of Lung Cancer: 2 cases in stage Ia, 3 cases in stage Ib, 3 cases in stage IIa, 1 case in stage IIb, and 1 case in stage IIIa. There were 3 cases in stage IIIb, 5 cases in stage IIIb, and 3 cases in stage IV. All 20 patients underwent chest enhanced computed tomography (CT) examination and found space-occupying lesions in the lungs. New organisms were seen in 7 patients under bronchoscopy and biopsied. One case was diagnosed as pulmonary LCNEC, the remaining 1 case was considered to have non-small cell lung cancer (NSCLC), 2 cases were considered to be SCLC, and 1 case was considered to have large lung cancer. Cell lung cancer, one case was considered to be severe dysplasia, and another case was considered to be chronic inflammation of the lung tissue. Two patients underwent lung puncture biopsy, of which one had mild aplasia and the other was prone to large cell lung cancer. Thirteen patients were diagnosed with pulmonary LCNEC by surgical pathology, and another 3 patients were diagnosed by superficial lymph node biopsy. Pathological diagnosis refers to the pathological diagnostic criteria proposed by WHO in 2004, including: ① Neuroendocrine morphology; ② High cell division ratio; ③ Necrosis; ④ NSCLC cytological characteristics; ⑤ Chromogranin A (CgA), synaptophysin, Syn) and CD56/neural cell adhesion molecule, at least one

of the three immunohistochemical markers was positive or neuroendocrine granules were visible under an electron microscope.

Computed tomography (CT) findings are mainly peripheral masses, which may be accompanied by uneven enhancement and necrosis. The observation sample set $D=\{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$ of a random variable can be regarded as obtained by independently sampling (experimenting) the same random variable N times [1]. It can also be seen as having N identical (same probability distribution) random variables X , each independently sampling to obtain a total of N observation samples.

We know that the occurrence probability of any sample x_i is, then the joint probability of all samples is $P(x_i, \theta) = P(x^{(1)}, x^{(2)}, \dots, x^{(N)}; \theta)$. And since all samples satisfy independent and identical distribution (i.i.d), the decomposition rule according to the joint probability distribution is

$$P(D; \theta) = P(X^{(1)}, X^{(2)}, \dots, X^{(N)}; \theta) = \prod_{i=1}^N P(x_i; \theta) \quad (4)$$

Assume that the possible value space of θ is Θ , denoted as $\theta \in \Theta$. No matter what value θ takes, there is a certain possibility (probability) to generate this sample set D , but obviously the value of θ will affect the probability of generating this sample $P(D; \theta)$ [2].

Maximum Likelihood Estimation

We obtain $\{\theta\}$ by maximizing the log-likelihood function $l(\theta; D)$, which is equivalent to maximizing the likelihood function $L(\theta; D)$ [3].

$$\log L(p) = (\sum x_i) \log(p) + (n - \sum x_i) \log(1 - p) \quad (5)$$

Although we take the spread and metastasis of normal cancer cells as an example of discrete random variables, maximum likelihood estimation can also be applied to parameter estimation of continuous-valued random variables [4]. Continuous-valued random variables use the probability density function to represent the probability of each state [5]. The probability density function $P(X=x; \theta)$ is used to replace $f(X=x; \theta)$. There is no need to solve the maximum likelihood function [6].

When the likelihood function of the given observation data reaches the maximum value, the value of the unknown parameter in the likelihood function is optimal. Since the spread and metastasis of lung cancer cells have already occurred, I assume that the probability of their occurrence is the greatest, and we believe that the parameter values that make these samples have the greatest probability of occurrence are the optimal values [7].

The log-likelihood function of the observation sample set is

$$\begin{aligned} l(\theta; D) &= \log L(\theta; D) \\ &= \log P(x^{(1)}, \dots, x^{(N)}; \theta) \\ &= \prod_{i=1}^N f(x_i; \theta) \\ &= \sum_{i=1}^N \log f(x_i; \theta) \end{aligned} \quad (6)$$

We then solve for the values of the parameters by maximizing the log-likelihood function. Of course, sometimes we can directly obtain the analytical solution by setting the partial derivative to 0. However, sometimes we cannot obtain the analytical solution and need to use gradient iteration to solve it. Although maximum likelihood estimation is widely used, it is not perfect. When the sample is small or the sample is biased, the estimated value obtained will have a larger deviation. We can use Bayesian estimation, which can be regarded as an upgraded version of maximum likelihood estimation, to solve this extreme scenario by adding a priori.

Bayesian Estimation

Parameter values are not fixed values but are uncertain because we do not observe them. For unobserved events, every value is possible [8]. In $P(X; \theta)$, the parameter θ should also be a variable with random value, so $P(X; \theta)$ should be the joint probability distribution $P(X; \theta)$, not the conditional probability distribution [9]. According to the chain rule, the joint probability can be decomposed into the product of conditional probabilities:

$$P(X; \theta) = P(\theta)P(X | \theta) \quad (7)$$

Among them, $P(\theta)$ is the probability distribution of variable θ , and $P(X|\theta)$ is the conditional probability distribution of X under the condition of known θ . At this time, the generation process of an observation sample of X is:

1. First obtain the sampling value $\bar{\theta}$ of θ from the probability distribution $P(\theta)$.
2. Then substitute $\bar{\theta}$ into the conditional probability distribution $P(X|\theta=\bar{\theta})$.
3. Finally, \bar{X} is obtained by sampling from the conditional probability distribution $P(X|\theta=\bar{\theta})$.

Variable x is the "cause" variable, variable X is the "effect" variable. Cause, it is the speed at which cancer cells spread. The effect is the patient's life time. We write the observation sample D of the variable and the variable θ in the form of Bayes' theorem:

$$P(\theta | D) = \frac{P(\theta | D)P'(\theta)}{P(D)} \quad (8)$$

$P(\theta|D)$ represents the probability distribution of the "causal" variable θ inferred based on the "result" and D . It is usually called the posterior probability distribution. Here "posterior" means the patient's life span [10]. The evidence

here It refers to the "observation results", that is, the observation sample set.

$P'(\theta)$ represents the empirical understanding of θ when there is no evidence (observation sample set), which is called the prior probability distribution [11]. The prior is generally based on specific data and empirically setting a common probability distribution. If you don't know anything about $P'(\theta)$, you can set it to a uniform distribution. $P(D|\theta)$ is the probability of generating an observation sample under the condition of θ . We know that the observation sample set is independent and identically distributed (i.i.d), so it has the following form after expansion:

$$P(D|\theta) = P(\{X(1), \dots, X(N)\} | \theta) = \prod_{i=1}^N P(X^{(i)} | \theta) \tag{9}$$

This is actually the likelihood of the sample, so $P(D|\theta)$ is the likelihood value of the sample.

Hypothetical test

Point estimators use sample statistics to estimate population parameters, but point estimators are not exact [12]. Different from point estimation and interval estimation, hypothesis testing is not used to estimate the population parameters, but to determine whether a certain hypothesis about the population (parameters) is true [13]. The process of hypothesis testing can generally be abstracted into four steps. We call the hypothesis about the population the null hypothesis, usually represented by the symbol H_0 . H_0 is a hypothesis or assertion about the population [14]. It is a virtual hypothesis. For example, in our example, the null hypothesis is: Assume that our maximum likelihood estimate is correct, that is, lung cancer patients die within 600 days due to the rapid spread of cancer cells. This is a dummy assumption we made, expressed symbolically as:

$$H_0 : \mu = 600$$

Null Hypothesis

In statistical significance testing, the assertion being tested is called the "null hypothesis." The test primarily assesses the sufficiency of the evidence that rejects the null hypothesis. The alternative hypothesis is usually represented by the symbol H_a . If the null hypothesis H_0 is not true, it means that the alternative hypothesis H_a is true, and usually the two are contradictory. In our statistics, the alternative hypothesis is:

$$H_0 : \mu \neq 600$$

The process of hypothesis testing is to first assume that H_0 is correct, and then look for evidence that negates H_0 under this premise. If "evidence" is found, and the evidence is strong enough, reject H_0 and accept H_a ; if there is not enough "Evidence", just accept H_0 . Hypothesis testing is simply a probabilistic selection of the most likely outcome.

Setting Decision Criteria

Hypothesis testing is to find "evidence" that negates H_0 . Usually this "evidence" is an "impossible" event that occurred under the condition that H_0 is established. The so-called "impossible event" is an event with a very low probability. The significance level is a standard for judging how small a low-probability event is. The smaller the significance level value, the smaller the probability of the event and the more extreme it is. Usually we use the symbol α to represent it. In hypothesis testing, if an event occurs with a probability less than or equal to α under the condition that H_0 is established, we will think that H_0 is likely to be wrong, and we will reject H_0 at this time.

Calculating the Test Statistic

After we have the test standard, we need to calculate a value and compare it with this standard, and calculate the statistics of this value [15]. Generally, we will select an appropriate test statistic based on the actual problem scenario, and then calculate the value of this test statistic and the theoretical possibility of obtaining this value. This probability value is called the P value. Finally, the P value is compared with the test standard value α , and a conclusion is given based on the comparison result.

The sampling distribution of the sample mean statistic \bar{X} is the normal distribution $N(\mu, \sigma^2/N)$. The probability of occurrence of the sample statistic value can be calculated through the sampling distribution of the sample statistic.

In our statistics, under the premise that H_0 is established, the mean (expected) time after metastasis of lung cancer cells is 600 days, and the overall variance is unknown. Temporarily represented by the symbol $\mu=600$, the sampling distribution of the mean statistic σ^2 of the sampling sample is

$$\bar{X} \sim N(\mu = 600, \sigma^2 / N) \tag{10}$$

The variance of the above sampling distribution $\text{Var}(\bar{X}) = \sigma^2/N$ is unknown, where σ^2 is the population variance parameter, N is the sampling sample size, usually the sample size N is known, and the sampling sample size is 20. At this time, it is also necessary to obtain the overall variance parameter σ^2 . We can use the sample variance to approximate the population variance:

$$s^2 = \frac{\sum_{i=1}^N (\bar{x} - x_i)^2}{N - 1} \tag{11}$$

Here we assume that the calculated population variance estimate is $\hat{\sigma}^2 = 36.0$, then the variance of the sampling distribution of the sample mean statistic is $\hat{\sigma}^2/N = 36.0$. Under the condition that H_0 is established, the sampling distribution of the sample mean statistic \bar{X} is:

$$\bar{X} \sim N(20, 0.36) \tag{12}$$

Theoretically the expectation of sample statistic \bar{X} is 20 and the variance is 0.36. Then we find that the sample mean calculated from the sample is $\bar{X}=20$. Theoretically, the closer the sample result is to 20, the more likely it is that H_0 is correct; the further the sample mean result deviates from 20, the more likely it is that H_0 is wrong.

Decision Making

The decision of hypothesis testing on the overall assertion is not 100% correct. The decision of accepting or rejecting the null hypothesis is based on probability, so we may make wrong decisions [16]. The significance level α is the upper limit of the probability of making a wrong decision. A decision is made based only on the statistical value of a sample, and the judgment of the decision is based on probability, so the conclusion given by the hypothesis test may also be wrong. There are four possible outcomes between the true situation of the null hypothesis and the test conclusion.

Table 1: Four decision outcomes of hypothesis testing

Item	Accept the null hypothesis	Reject the null hypothesis
Null hypothesis is true	Correct $1-\alpha$	Type I error β
The null hypothesis is false	Type II error β	Correct $1-\alpha$

Table 1 presents four situations in tabular form, two of which are correct and the other two are incorrect.

If the null hypothesis is true and the decision result is accepted, then the decision result is correct, and the probability of this result is $1-\alpha$.

The null hypothesis is true, but the decision result is rejection. At this time, the decision result is wrong. The probability of this result is α .

The null hypothesis is false, but the decision result is accepted. At this time, the decision result is wrong. The probability of this result is β . This is called a Type I error.

If the null hypothesis is false and the decision result is rejection, the decision result is correct. The probability of this result is $1-\beta$. This is called a Type II error.

Conclusion

Although various prediction models can help improve clinical predictability, prediction results cannot be understood blindly or in extremes. Each patient is unique, and we can only observe but not determine his or her final survival time.

Talking about survival time with cancer patients and their families is a very heavy and contradictory matter. We should not argue against the value of predicting survival, but neither should we exaggerate the accuracy of prediction results. The

spread of cancer does not mean the immediate end of life or the end of dreams. If it can extend the life cycle by 1 year, 3 years, 5 years, or longer. Medicine may have more and newer treatment methods and may achieve better control. It is also a blessing to turn cancer into a chronic disease. Taking advanced non-small cell lung cancer as an example, immunotherapy represented by PD-1/PD-L1 inhibitors has greatly improved the survival benefits of patients.

A recent study published in the "Journal of Clinical Oncology" showed that K drug (pembrolizumab) is used in the first-line treatment of PD-L1 TPS (the percentage of PD-L1 stained tumor cells in tumor cells) $\geq 50\%$, EGFR /ALK mutation-negative non-small cell lung cancer patients, the three-year overall survival rate was 43.7%, the four-year overall survival rate was 35.8%, and the five-year overall survival rate also reached 31.9%.

Consent to Participate

Additional informed consent was obtained from all individual participants for whom identifying information is included in this article.

Funding Declaration

The authors did not receive support from any organization for the submitted work. No funding was received to assist with the preparation of this manuscript. No funding was received for conducting this study.

Clinical Trial Number

Clinical trial number: not applicable.

Ethics declarations

Competing interests

The authors declare no competing interests.

Ethical Approval

Human Ethics and Consent to Participate declarations: not applicable

Conflict of Interest

The authors declare no conflict of interest.

Consent to Publish Declaration

Not applicable

References

- Alves J M, Prado-Lopez S, Cameselle-Teijeiro J M, et al. Rapid evolution and biogeographic spread in a colorectal cancer. Nature communications 10 (2019): 5139.
- Bellomo D, Arias-Mejias S M, Ramana C, et al. Model combining tumor molecular and clinicopathologic risk

- factors predicts sentinel lymph node metastasis in primary cutaneous melanoma. *JCO precision oncology* 4 (2020): 319-334.
3. Vaghi C, Rodallec A, Fanciullino R, et al. Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors. *PLoS computational biology* 16 (2020): e1007178.
 4. Ryser M D, Min B H, Siegmund K D, et al. Spatial mutation patterns as markers of early colorectal tumor cell mobility. *Proceedings of the National Academy of Sciences* 115 (2018): 5774-5779.
 5. Chan J M, Quintanal-Villalonga A, Gao V R, et al. Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer cell* 39 (2021): 1479-1496.
 6. Hu Z, Li Z, Ma Z, et al. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nature genetics* 52 (2020): 701-708.
 7. Kim I G, Hu X G, Wang H J, et al. MiR-182-5p knockdown targeting PTEN inhibits cell proliferation and invasion of breast cancer cells. *Yonsei Medical Journal* 60 (2019): 148-157.
 8. Alves J M, Prado-Lopez S, Cameselle-Teijeiro J M, et al. Rapid evolution and biogeographic spread in a colorectal cancer. *Nature communications* 10 (2019): 5139.
 9. Boodaghi M, Libring S, Solorio L, et al. A Bayesian approach to estimate the diffusion coefficient of Rhodamine 6G in breast cancer spheroids. *Journal of Controlled Release* 340 (2021): 60-71.
 10. Alharbi Y S, and Kamel A. Fuzzy System Reliability Analysis for Kumaraswamy Distribution: Bayesian and Non-Bayesian Estimation with Simulation and an Application on Cancer Data Set. *WSEAS Transactions on Biology and Biomedicine* 19 (2022): 118-139.
 11. Chroni A, Miura S, Oladeinde O, Aly V, et al. Migrations of cancer cells through the lens of phylogenetic biogeography. *Scientific Reports* 11 (2021): 17184.
 12. Liu Y, Edrisi M, Ogilvie H A, et al. NestedBD: Bayesian inference of phylogenetic trees from single-cell DNA copy number profile data under a birth-death model. *bioRxiv* (2022).
 13. Münch M M, van de Wiel M A, van der Vaart A W, et al. Semi-supervised empirical Bayes group-regularized factor regression. *Biometrical Journal* 64 (2022): 1289-1306.
 14. Pouymayou B, Balermipas P, Riesterer O, et al. A Bayesian network model of lymphatic tumor progression for personalized elective CTV definition in head and neck cancers. *Physics in Medicine & Biology* 64 (2019): 165003.
 15. Kumar S, Chroni A, Tamura K, et al. PathFinder: Bayesian inference of clone migration histories in cancer. *Bioinformatics* 36 (2020): 675-683.
 16. Susswein Z, Sengupta S, Clarke R, et al. Borrowing ecological theory to infer interactions between sensitive and resistant breast cancer cell populations. *BioRxiv* (2022).



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC-BY\) license 4.0](https://creativecommons.org/licenses/by/4.0/)