





ABSORBING MARKOV CHAINS TO CHARACTERIZE AND PREDICT METASTASIS PATHWAYS IN CHILDHOOD CANCER

DAVID H. MARGARIT ^{*,†,¶}, MARCELA V. REALE ^{*,‡},
ARIEL F. SCAGLIOTTI ^{‡,§} and LILIA M. ROMANELLI ^{*,†}

^{*}*Instituto de Ciencias (ICI)*

Universidad Nacional de General Sarmiento (UNGS)
J. M. Gutiérrez 1150, Los Polvorines

Malvinas Argentinas B1613, Buenos Aires, Argentina

[†]*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*
Godoy Cruz 2290, Ciudad Autónoma de Buenos Aires C1425, Argentina

[‡]*Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT)*
Universidad Nacional de La Matanza (UNLaM)

Florencio Varela 1903, San Justo, La Matanza B1754, Buenos Aires, Argentina

[§]*Facultad Regional Mendoza (FRM), Universidad Tecnológica Nacional (UTN)*
Coronel Rodríguez 273, Ciudad de Mendoza M5500, Mendoza, Argentina

[¶]*dmargarit@campus.ungs.edu.ar*

Received 28 April 2023

Accepted 22 August 2023

Published

Cancer and its metastasis in children have a high degree of lethality and side effects, so its characteristics need to be studied independently of adult cancer. The aim of this work is to model the metastasis pathways of the main childhood cancers worldwide by Absorbing Markov Chains, an important mathematical tool used for different applications in science. Statistical information was collected to detect the main affected organs (primary sites) and those where cancer cells generally spread and metastasize (secondary sites). Taking into account that it is a branching process, a directed graph was developed, and the associated transition matrices for the first and second metastases were constructed. Organs whose cancers generally remain encapsulated and do not spread their cancer cells are considered absorbing states in terms of Markov processes. For the selected organs, the probability of ending up in each of the absorption states according to the primary site was calculated, as well as the number of possible previous metastases until reaching one of these states. Although the lung in childhood cancer is not a characteristic primary site, it is one of the main sites of metastasis. Therefore, this work dedicates a section to including this organ as a site of possible metastasis.

Keywords: Stochastic Analysis; Absorbing Markov Chains; Biological Modeling; Childhood Cancer; Metastasis Routes.

[¶]Corresponding author.

2 *Margarit et al.*

1. Introduction

Cancer is the deadliest disease in the world, according to the World Health Organisation (WHO). One in six deaths is due to this disease that afflicts people at any age (children, adults, and the elderly).¹ Their causes are very varied.² In general, cancer is the uncontrolled replication of cells in a tissue, altering its intrinsic properties and damaging it, even causing death.³ Cancer can not only establish itself in the organ where it originates, but it can also spread and colonize other tissues or organs, generating a new one there. This phenomenon is known as metastasis.³ Although it is estimated from and to which organ metastasis can be generated, it is unknown what these routes will be like when there is subsequent cancer (metastasis of metastasis) or if cancer ends up in a tissue and does not spread further.³ The spread of metastasis is the leading cause of death from cancer because of its branched nature, in which various organs are involved, reducing life expectancy.

Health in children is seriously affected by cancer, causing premature deaths or serious side effects in those who manage to recover.⁴ It cannot be ignored, and it is part of the motivation for this work, that the most common pediatric cancers are not exactly the same for adults. This is because, in the case of adults, they are promoted for harmful habits (for example, smoking or unhealthy diets).^{5,6} As well, the sites of metastasis or the frequency with which they occur are not similar.⁵

When looking at metastasis pathways, models typically focus on one organ and its compatibility with spreading cancer cells to others. These models are diverse, from how cancer cells migrate through the circulatory system using fluid dynamics⁷ to an integrative genomic analysis (mutation burden, gene mutation frequency, mutation signature, etc.) of nasopharynx cancer.⁸ In the case of stochastic and computational models, Xie *et al.*⁹ reviewed different inferential, computational, and mathematical approaches to modeling cancer evolution; in Ref. 10, a mathematical model for chemoimmunotherapy treatment was developed; in Ref. 11, the routes of spread of metastases through lymphatics and blood vessels are analyzed using a Bayesian biogeographical approach; in Ref. 12, used Markov Chains to study metastases from bone to the chest wall, lung, or liver; and, in Ref. 13, also using Markov Chains, established probabilities of life expectancy for patients with breast cancer at different stages of tumor evolution.

Thus, from a mathematical point of view, Markov Chains promote a tool that is used in different fields due to its potential for predictability based on probabilities that are determined by linear algebra. Markov Chains are used not only in mathematics but also in other sciences.^{14–16} In this sense, it is possible to extrapolate and use them in biological systems, in particular in cancer and metastasis.

In previous work, we used Markov Chains for metastasis site research for adults in Argentina.¹⁷ Now, taking into account the importance of modeling mathematically the routes of metastasis in childhood cancer, we analyze metastatic pathways in the main hematological and solid cancers for children aged 0–14 years. The data

used here was obtained from the International Agency for Research on Cancer (IARC),¹⁸ belonging to WHO (more details in Sec. 2).

In this way, by using Markov Chains, we will analyze the different steps between the metastases from one organ (the primary site) to others, whether they are secondary or tertiary sites. Those organs with a low probability of generating metastases are called absorbing states; therefore, to contemplate this phenomenon, this work is developed through Absorbing Markov Chains. Thus, we will probabilistically analyze how organs are related to childhood metastases and the role of absorbing states through current information. This analysis will contemplate two scenarios: the first one where lung cancer is not considered and the second scenario where it is. Even though lung cancer is not the most common pediatric cancer, this will be done in this way because it is reported to be an important metastasis site.

This paper is organized as follows. In Sec. 2, we introduce the way in which the data was obtained, the criteria to rank them, the tools to be used, and the graph on which we base ourselves for the development of the investigation. In Sec. 3, we detail how the transition matrix is constructed and the discrimination by type of state (absorbent and non-absorbent), characterizing the metastasis route for the first generation of the same without taking into account the lung as a possible state. The analysis of the second step of metastases is described in Sec. 4. The absorptive capacity of non-absorbing states, those with almost zero probability of generating metastasis compared to the others analyzed, and the number of expected metastases are studied in Sec. 5. Next, in Sec. 6, we make an analysis similar to that developed in the previous sections but with the inclusion of the lung as a possible site of metastasis, discussing the differences and/or coincidences with the case where it was not taken into account. Finally, Sec. 7 concludes.

2. Methodology

There are three general ways to spread cancer cells and generate new cancer in another organ: through the blood circulation, lymph nodes, and transcoelomic tissues.¹⁹ Blood vessels are the primary route for spreading to distant organs. Meanwhile, lymphatic vessels provide a route to local lymph nodes, and frequently, after metastasis, they spread through the blood. Although spreading through the blood seems to be quite common, the lymphatic spread appears to depend on the location of the primary tumor. For example, cancers of bone and soft tissue (sarcomas) first spread through the blood, while kidney cancer spreads through the lymphatic system. Transcoelomic spread is rare and appears to be restricted to mesotheliomas and ovarian carcinomas.²⁰

The information obtained for the main cancers from the current WHO-IARC statistics was ordered by the highest number of cases worldwide and with a crude rate equal to or greater than 0.6 in the period 2010–2020. The crude rate is, for each specific cancer, calculated by dividing the number of new cancers observed by every 100,000 people at risk in a year.²¹ There are two types of cancer that we

4 *Margarit et al.*

have arbitrarily added to the main ones, which are bone cancer and lung cancer. Despite bone cancer not being a main primary site, it is one of the main sites where other childhood cancers migrate.^{22–26} In the same sense, the lung is not a common primary site, but it is one of the most important sites where cancer cells can migrate (colonizing this organ) and from there do the same in others.^{24,27–32} The lung, both in children and adults, is one of the main metastasis links in the human body.^{17,19,33–35}

Those cancers that commonly affect children are related to the circulatory and lymphatic systems (leukemia, Hodgkin’s lymphoma, and non-Hodgkin’s lymphoma).¹⁸ Subsequently, solid tumors are the ones that continue the list (brain, kidney, liver, ovary, testicular, thyroid, nasopharynx, lip and oral cavity, lung, and bone).¹⁸ However, bone, liver, and brain are organs that, when they are primary sites, do not generate metastases in other organs (or at least there is a very low probability of this happening)^{18,19}; these organs are called absorbing states or sites.

It is worth clarifying that we will not distinguish between cancer subtypes with the intention of being clearer and more global in our analysis. The most common secondary sites for each cancer were obtained from a literature search. The latter is very extensive and varied, which shows a need to develop a unified database with the most frequent secondary sites depending on the primary one. Thus, different investigations that contribute to this problem would be strengthened. Then, from the most reported cancers in children by WHO-IARC statistics,¹⁸ we found that these primary sites can generate metastasis; these sites (secondary sites) are depicted in Table 1.

Based on the data collected in Table 1, the directed graph of Fig. 1 was constructed. The directionality (or bi-directionality) of the arrows indicates where cancer cells can migrate from the primary site and create new cancer in one or more organs (secondary sites). The cases where the arrows leave and arrive at the same node (Br, Bo, and Li) indicate that they do not generate cancer in other organs (absorbing sites). In general, brain, liver, and bone cancers do not spread to other organs. Therefore, if they generate metastasis, it is only on themselves.^{45,46}

3. Construction of the Transition Matrix and Absorbing States Without Lung Cancer

As mentioned above, we will first consider the case where the lung is not included. Figure 1 shows the directed graph that will be the basis for the development of the formalism from Markov Chains.

Let X_0 (primary site) be the organ where cancer originated, and X_1 is the state of the process where a new cancer is formed (secondary site/s) coming from a primary site. In other words, X_0 refers to cancer generated in any of the organs (or nodes) that appear in Fig. 1 (or the first column of Table 1). X_1 is the state where a secondary site is colonized by cancer cells from a specific primary site. The

Table 1. Table of main secondary sites for the most common childhood cancers.

| Primary site | Symbol | Type | Main secondary sites | References |
|----------------------|--------|---------------|--|------------------------|
| Leukemia | Le | Hematological | Ovaries, testicles, and brain | Refs. 36-38 |
| Hodgkin lymphoma | H | Lymphatic | Testicles and brain | Refs. 37, 39 and 40 |
| Non-Hodgkin lymphoma | NH | Lymphatic | Testicles and brain | Refs. 37, 39 and 40 |
| Kidney | K | Solid | Brain, liver, bone, and lung | Refs. 5, 22 and 41 |
| Ovaries | O | Solid | Liver, bone, and lung | Refs. 23, 31 and 42 |
| Testicles | Te | Solid | Brain, liver, bone, and lung | Refs. 5, 24, 30 and 32 |
| Thyroids | Th | Solid | Brain, liver, bone, and lung | Refs. 25, 43 and 44 |
| Nasopharynx | N | Solid | Lip/oral cavity, liver, and lung | Refs. 27 and 28 |
| Lip/oral cavity | LO | Solid | Nasopharynx, bone, and lung | Refs. 26 and 29 |
| Lung | Lu | Solid | Kidney, ovaries, testicles, thyroid, nasopharynx, and lip/oral | Refs. 24 and 28-32 |
| Brain | Br | Solid | Does not usually metastasize to other organs | Refs. 45 and 46 |
| Liver | Li | Solid | Does not usually metastasize to other organs | Refs. 45 and 46 |
| Bone | Bo | Solid | Does not usually metastasize to other organs | Refs. 45 and 46 |

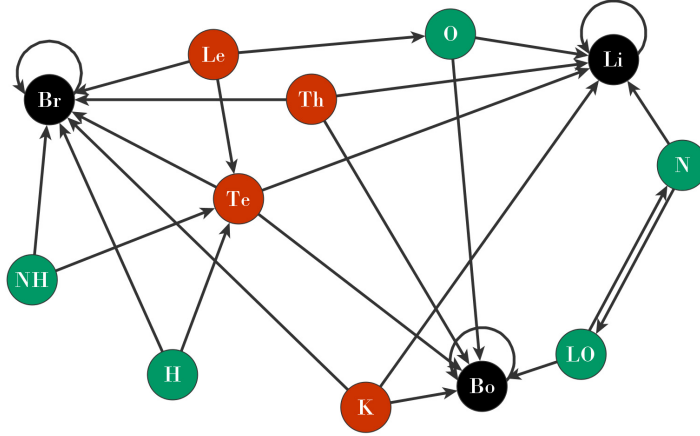
6 *Margarit et al.*

Fig. 1. Graph of metastasis routes, only from primary to secondary sites, for the most common childhood cancers (excluding lung), based on information collected in Table 1. Nodes in green: metastasis is generated from these organs mainly in 2 organs; in red: these organs generate metastases mainly in 3 organs; finally, in black, the absorbing states, these generally do not metastasize.

secondary sites are the nodes where the arrows finish in the graph (in the fourth column of Table 1 are the possible secondary sites for each organ).

We consider it a step in the Markov matrix when a new tumor has already been found in another organ. Likewise, the same concept applies to the second step matrix (this will be discussed later). In terms of the Markov Chain formalism,⁴⁷ the above implies that the transition matrix has been built under the assumption of the existence of the tumor in X_0 that will develop metastasis in X_1 . Then, the probability that an organ (i) develops metastasis into another (j) is

$$p_{ij} = P[X_1 = j \mid X_0 = i], \quad (3.1)$$

where $i, j = 1, 2, 3, \dots, m$ refer to the $m = 12$ (or 13 if considering the lung) organs taken into account in Table 1.

In other words, X_1 (metastasis from the primary site or a secondary site) will have a probability of being metastasis from some primary site called X_0 . If the tertiary site (metastasis from the secondary site) exists, it will be labeled as X_2 .

Generalizing these probabilities p_{ij} ($p_{ij} \geq 0$) for all the m cancers, we construct the probability transition matrix P .

$$P = [p_{ij}] = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1m} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2m} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & p_{m3} & \cdots & p_{mm} \end{pmatrix}, \quad \text{with } P \in \mathbb{R}^{m \times m}.$$

Since there is no general agreement on the probability of metastasis in children, because this differs in the reported cases, we will consider that the sites of metastasis are equiprobable. That is, if an organ i has ρ possible metastases, the probability p_{ij} will be $\frac{1}{\rho}$.

In terms of the current statistics and data, we built the probability transition matrix P as follows:

$$P = \begin{matrix} & \begin{matrix} \text{Le} & \text{H} & \text{NH} & \text{K} & \text{O} & \text{Te} & \text{Th} & \text{N} & \text{LO} & \text{Br} & \text{Li} & \text{Bo} \end{matrix} \\ \begin{matrix} \text{Le} \\ \text{H} \\ \text{NH} \\ \text{K} \\ \text{O} \\ \text{Te} \\ \text{Th} \\ \text{N} \\ \text{LO} \\ \text{Br} \\ \text{Li} \\ \text{Bo} \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

A way to extract information from P : in Figs. 2(a) and 2(b) there are two examples that describe the probabilities of generating metastasis in determinate organs from specific primary sites (thyroids, lip, and oral cavity).

We can observe in the matrix P the existence of Absorbing States in the system. When these states are reached, it is impossible to leave.⁴⁷ In terms of probability:

$$p_{ii} = 1 \quad \text{and} \quad p_{ij} = 0. \quad (3.2)$$

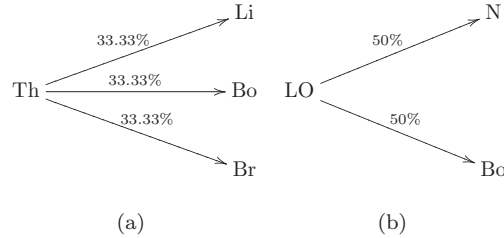
8 *Margarit et al.*

Fig. 2. Probabilities of thyroids (a) and lip/oral cavity (b) of generating metastasis given by matrix P .

In our system, we can see that in the matrix P , there are three absorbing states (the brain, liver, and bone). In these particular organs, when cancer originates there, the probability of metastases in another organ is virtually null. However, if the cancer forms elsewhere (first or second), it can metastasize to one of these sites, but no new metastases will form from there. Therefore, it can be considered an absorbing matrix.

4. Transition Matrix for Tertiary Sites

A quite important application of Markov Chains is to analyze the probability when a process goes from state i through an intermediate state k to state j in two steps.⁴⁷ That is

$$p_{ij}^2 = P[X_2 = j \mid X_0 = i] = \sum_{k=1}^m p_{ik} \cdot p_{kj}. \quad (4.1)$$

In the context of our models, we can quantitatively estimate the probability of metastasis being generated in a tertiary site, that is, a metastasis of metastases. The intermediate state k can be a metastasis from potential organs where there is already a metastasis generated from the primary site i , consequently, we lose precise information about which is the intermediary organ or tissue. Generalizing this for all organs, based on Eq. (4.1), the matrix of probabilities of metastases from metastases (cancer to the tertiary site from the primary site) is given by

$$P^2 = [p_{ij}^2] = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1m} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2m} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & p_{m3} & \cdots & p_{mm} \end{pmatrix}^2, \quad \text{with } P^2 \in \mathbb{R}^{m \times m},$$

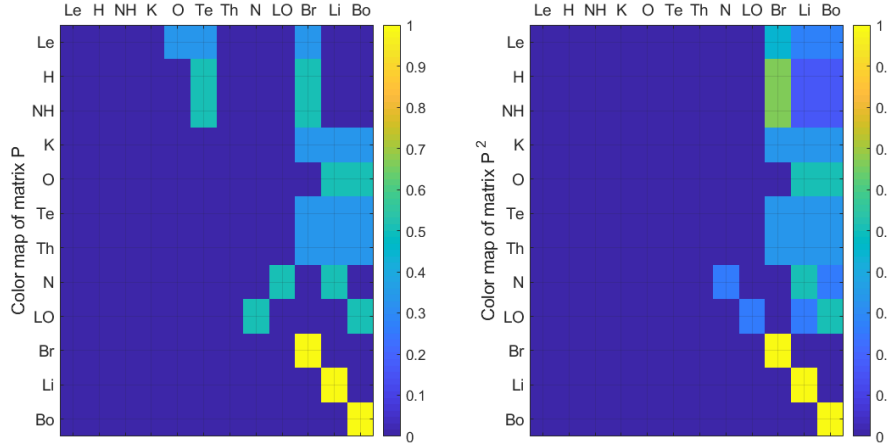


Fig. 3. Color map for transition matrices for the first (P) and second (P^2) step of the metastasis network without lung.

where with the expression mentioned above, we obtain

$$P^2 = \begin{pmatrix} \text{Le} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{9} & \frac{5}{18} & \frac{5}{18} \end{pmatrix} \\ \text{H} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \\ \text{NH} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \\ \text{K} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \\ \text{O} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ \text{Te} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \\ \text{Th} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \\ \text{N} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{4} \end{pmatrix} \\ \text{LO} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \\ \text{Br} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \\ \text{Li} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \\ \text{Bo} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{pmatrix}.$$

The transition matrix of the second step (P^2) shows that, in general, from a primary site there is a higher probability, compared to P , of having reached an absorbing state (liver, liver, or brain). That is, for metastases of metastases, an absorbing state always has a greater than zero probability of being reached. Another way of showing the transition matrices P and P^2 is through a color map for each of them, as we can see in Fig. 3. This is with the intention of visually highlighting the

10 *Margarit et al.*

probabilities in each of the matrices and seeing the impact on the transition when passing from the first to the second step (from the first to the second metastasis). Going from blue to yellow, blue implies a probability of zero at that position in the matrix (remember that they are exactly the same as at P and P^2) and yellow implies a probability close to 1 that metastasis will be generated there. To quantify this information, such as the number of steps to complete an absorption state, both in general and for a particular organ, we will develop the following section: It is important to note that when we talk about cancer spread or metastasis, we do so for simplistic reasons since the states of the Markovian process X_0 , X_1 and X_2 were previously defined in this work. We must emphasize that what is really disseminated are cancer cells from a primary site, which circulate and eventually colonize another organ (secondary sites).

5. Absorptive Capacity by Non-Absorbing States and Number of Expected Metastases

In Absorbing Markov Chains, the number of steps before the system is absorbed and the probability of absorption by any absorbing state can be found. In order to calculate this information, each transition matrix must be represented in its canonical form,⁴⁷ called J . The canonical matrix J is composed by 4 sub-matrices (N , A , 0 , and I_a). These smaller matrices contain elements of probability that originate a absorbing states and n non-absorbing states. Thus, there are $a + n = m$ states of the system. For our model, the matrix J is represented by

$$J = \begin{array}{c} \text{Le} \\ \text{H} \\ \text{NH} \\ \text{K} \\ \text{O} \\ \text{Te} \\ \text{Th} \\ \text{N} \\ \text{LO} \\ \hline \text{Br} \\ \text{Li} \\ \text{Bo} \end{array} \left(\begin{array}{c|c} \text{Le H NH K O Te Th N LO} & \text{Br Li Bo} \\ \hline & \\ & \\ & \\ & \\ & \\ & \\ & \\ & \\ & \\ \hline & \\ & \\ & \end{array} \right) \begin{array}{c} \\ \\ \\ \\ \\ N \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ 0 \\ \\ I_a \end{array} \cdot$$

- Sub-Matrix N : $N \in \mathbb{R}^{n \times n}$ has the probabilities of going from a non-absorbing state to another non-absorbing state.

- Sub-Matrix A : $A \in \mathbb{R}^{n \times a}$ has the probabilities of going from a non-absorbing state to another absorbing state.
- Sub-Matrix 0 : $0 \in \mathbb{R}^{a \times n}$ represents the probabilities of going from an absorbing state to another non-absorbing state (zeros matrix).
- Sub-Matrix I_a : $I_a \in \mathbb{R}^{a \times a}$ this represents the probabilities of remaining inside of an absorbing state (an identity matrix).

Let be the fundamental matrix⁴⁷ $F = (I - N)^{-1}$, where I is an identity matrix with the same dimensions as N . Starting in a transient state, the expected number of steps T (in our case, a step in the metastasis propagation from one organ to another) before being absorbed by an absorbing state, T , is given by

$$T = F \cdot c, \quad (5.1)$$

where c is a column vector whose entries are ones.

$$T = \begin{matrix} & \text{steps} \\ \text{Le} & \left(\begin{array}{c} 5 \\ 3 \end{array} \right) \\ \text{H} & \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \\ \text{NH} & \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \\ \text{K} & 1 \\ \text{O} & 1 \\ \text{Te} & 1 \\ \text{Th} & 1 \\ \text{N} & 2 \\ \text{LO} & 2 \end{matrix} .$$

As can be seen, the expected number of steps T for metastasis generated in a primary site and ending in an absorbing organ is at most (or close to) 2. Hematological or lymphatic cancers have a number of steps between 1 and 2, which implies that more than one metastasis can occur from the secondary site. Although these cancers do not metastasize in adults (except for very isolated events), in children they can develop a branching metastatic pathway. In the cases of the nasopharynx and lip/oral cavity, they are generally absorbed in the second step of metastasis until they reach one of the three absorbing organs. Finally, ovarian, testicular, and thyroid cancer cells will migrate and metastasize but remain in an absorbent state, improving the prognosis for prevention.

The vector T can be interpreted as a degree of tolerance in the human body. Generally, if the primary sites are not treated, they are known to trigger eventual metastases (in one or more organs) and from there to other sites. Therefore, this

12 *Margarit et al.*

information acquired by T is an important tool for early diagnosis and potential predictions about cancer sites and pathways, depending on their origin.

On the other hand, the probability of absorption of any non-absorbing state by any absorbing state is depicted by the matrix Z^{47} :

$$Z = ((I - N)^{-1}) \cdot A, \quad (5.2)$$

which means

$$Z = \begin{matrix} & \text{Br} & \text{Li} & \text{Bo} \\ \text{Le} & \left(\frac{4}{9} & \frac{5}{18} & \frac{5}{18} \right) \\ \text{H} & \left(\frac{2}{3} & \frac{1}{6} & \frac{1}{6} \right) \\ \text{NH} & \left(\frac{2}{3} & \frac{1}{6} & \frac{1}{6} \right) \\ \text{K} & \left(\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \right) \\ \text{O} & \left(0 & \frac{1}{2} & \frac{1}{2} \right) \\ \text{Te} & \left(\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \right) \\ \text{Th} & \left(\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \right) \\ \text{N} & \left(0 & \frac{2}{3} & \frac{1}{3} \right) \\ \text{LO} & \left(0 & \frac{1}{3} & \frac{2}{3} \right) \end{matrix}.$$

The matrix Z has a similar meaning as the vector T , except that we observe specifically for each primary site the probability of being absorbed by a certain absorbing organ (states). These results also provide sensitive information for the treatment of childhood cancer. As can be seen from the matrix Z , leukemia and both lymphomas are more likely to reach (if they metastasize) the brain. In the case of the nasopharynx and lip/oral cavity, they finish their route in the liver or

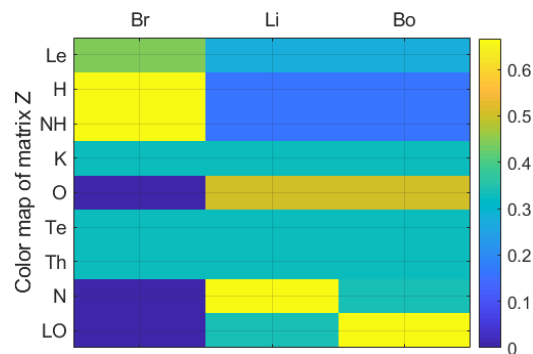


Fig. 4. Color map of matrix Z : probability of absorption of any non-absorbing state by absorbing states (network without lung).

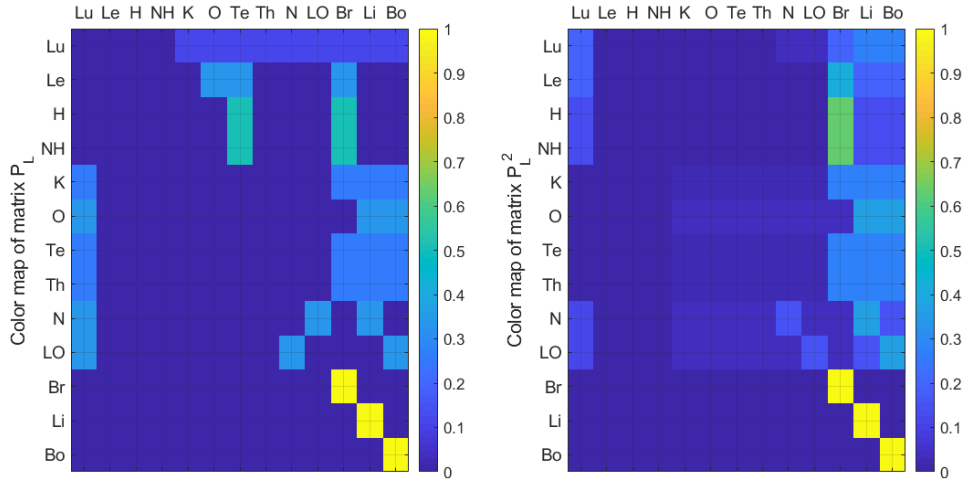


Fig. 6. Color map for transition matrices for the first (P_L) and second (P_L^2) step for metastasis network with lung.

Here, P_L^2 shows that organs like the lung, kidney, ovaries, testicles, and thyroid are also possible sites of metastases, besides those that already appeared in P^2 from Sec. 4 (lip/oral cavity, nasopharynx, brain, liver, and bone). Therefore, the lung is one of the most important nexuses between the primary and tertiary sites. An alternative method, mainly for ease of visualization, is to view the information provided by P_L and P_L^2 by using a color map (Fig. 6). Remember that the scale of this color map is from blue to yellow (0 to 1).

The expected number of steps T_L is

$$T_L = \begin{matrix} & \text{steps} \\ \begin{matrix} \text{Lu} \\ \text{Le} \\ \text{H} \\ \text{NH} \\ \text{K} \\ \text{O} \\ \text{Te} \\ \text{Th} \\ \text{N} \\ \text{LO} \end{matrix} & \begin{pmatrix} 2.31 \\ 2.11 \\ 1.79 \\ 1.79 \\ 1.58 \\ 1.77 \\ 1.58 \\ 1.58 \\ 2.65 \\ 2.65 \end{pmatrix} \end{matrix}.$$

Here, when T and T_L are compared, it is possible to observe that the values of T_L are slightly higher, but this does not make a big difference: On average, the

16 *Margarit et al.*

expected number of steps for metastasis generated in a primary site and ending in an absorbing organ is 1.981, that is 2.

$$Z_L = \begin{matrix} & \begin{matrix} \text{Br} & \text{Li} & \text{Bo} \end{matrix} \\ \begin{matrix} \text{Lu} \\ \text{Le} \\ \text{H} \\ \text{NH} \\ \text{K} \\ \text{O} \\ \text{Te} \\ \text{Th} \\ \text{N} \\ \text{LO} \end{matrix} & \left(\begin{array}{ccc} \frac{21}{83} & \frac{31}{83} & \frac{31}{83} \\ \frac{116}{249} & \frac{133}{498} & \frac{133}{498} \\ \frac{109}{166} & \frac{57}{332} & \frac{57}{332} \\ \frac{109}{166} & \frac{57}{332} & \frac{57}{332} \\ \frac{26}{83} & \frac{57}{166} & \frac{57}{166} \\ \frac{7}{83} & \frac{38}{83} & \frac{38}{83} \\ \frac{26}{83} & \frac{57}{166} & \frac{57}{166} \\ \frac{26}{83} & \frac{57}{166} & \frac{57}{166} \\ \frac{21}{166} & \frac{373}{664} & \frac{207}{664} \\ \frac{21}{166} & \frac{207}{664} & \frac{373}{664} \end{array} \right) \end{matrix}.$$

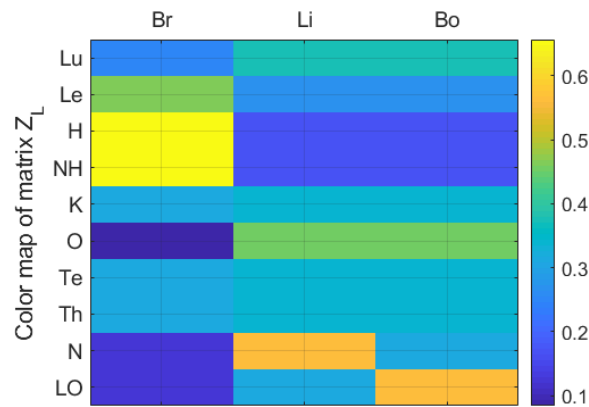


Fig. 7. Color map of matrix Z_L : probability of absorption of any non-absorbing state by absorbing states (network with lung).

By analyzing Z_L , it can be seen that there is a small probability that the ovaries, nasopharynx, and lip/oral cavity finish their metastasis route in the brain, unlike when the lung is not considered. Therefore, all organs end up being absorbed by any of the three absorbing states. In biological terms, this means that sooner or later, cancer cells will colonize one of these organs (as long as the organism tolerates such a degree of metastatic spread in the organs). Figure 7 shows the matrix Z_L (in the same way as was made with Z) as a color map with a scale between 0 (blue) and 1 (yellow), in order to depict in an easy way each probability by organ, but considering the lung in the network.

7. Conclusions

Solutions to the problem of cancer cannot be thought of exclusively as palliative or curative therapies when there are physical and mathematical mechanisms that can probabilistically quantify the sites where different types of malignant tumors can spread. Clearly, when the diagnosis is late (by this we mean that cancer has already been diagnosed in a primary site), it is important and urgent to determine the places where it can spread. This, as we already mentioned in the introduction to this paper, contributes significantly to more efficient and earlier therapies for the fight against cancer and its spread in the body, particularly in children, where the aggressiveness of this disease is greater. Thus, the results of this work add relevance to this field of interdisciplinary research.

Through the use of Absorbent Markov Chains, we were able to quantitatively characterize the metastatic pathways of the main childhood cancers, whether solid, lymphatic, or hematological, in two cases: when the lung is not considered and when it is.

It is important to highlight that it was found that the absorption states (liver, brain, and bone) are reached in two steps on average. In addition, the probabilities of knowing specifically in which absorptive organ the pathway ends, depending on the primary site, were characterized.

In both cases, whether considering the lung or not, the most important (in terms of probability) possible sites of metastasis are the lips/oral cavity, nasopharynx, brain, liver, and bones. Particularly in the case where the lung is taken into account, more possible sites are added: kidney, ovaries, testicles, thyroid, and lung. The main difference, in this case, is that these probabilities are lower (at most a quarter) than those of the other five organs. Thus, depending on the degree of depth with which each case will be analyzed, it will be convenient to use one or the other of these methods of research. The matrices T , T_L , Z , and Z_L provide relevant information for any preventive treatment in children experiencing this disease.

The spread of cancer cells from a primary site and their colonization of other nearby or distant organs (secondary sites) may be simultaneous. This means that when we analyze what happens with a particular organ, other chains from other organs can happen in parallel. In the same way, when we talk about absorption

probabilities by an absorbing organ (Br, Bo, and Li), we do it for a certain chain, that is, a specific organ. Eventually, a branching process in some of the intermediate organs may be triggered until they fall into an absorbing state. The latter is not contemplated in the model, but it is worth mentioning. However, this model is robust and gives a relevant approximation of what happens with the routes of metastasis of the main organs in childhood cancer.


In this work, we have assumed the equiprobability of cancer spreading to other possible sites. This is a generalization that is due to the lack of information on organ metastasis rates in childhood cancer. This does not happen in adults; there is information from autopsies,^{48,49} where metastasis rates are quantified according to the primary site. Although this information, through autopsies, is not available for children (we can assume that it is due to the extreme sensitivity that such a process implies). In the same way, the data about metastasis in children is in very low populations to make statistics and a distribution by organ. It will take future work to delve into the subjects in order to optimize the model developed with unequal transition probabilities.

Finally, and continuing with the previous paragraph, future efforts will be devoted to expanding the criteria of this work. This can be done by considering different ethereal ranges and comparing them with the intention of finding significant patterns. Also, implement this model to discriminate by sex or different pathologies related to some types of cancer in particular. Also, implementing machine learning to automatically optimize the model when the available statistics are updated would be a noticeable improvement.

Acknowledgments


This work is financed by PIP-CONICET No. 11220200100439CO and the National Agency for the Promotion of Research, Technological Development, and Innovation PICTO-2021-UNGS-00010.

ORCID

David H. Margarit  <https://orcid.org/0000-0003-1946-0413>

Marcela V. Reale  <https://orcid.org/0000-0002-9856-7501>

Ariel F. Scagliotti  <https://orcid.org/0000-0002-8730-3885>

Lilia M. Romanelli  <https://orcid.org/0000-0002-9272-6575>

References

1. World Health Organization International Agency for Research on Cancer, 2023, <https://www.who.int/news-room/fact-sheets/detail/cancer>.
2. Blackadar CB, Historical review of the causes of cancer, *World J Clin Oncol* **7**(1):54, 2016.

3. Hiatt HH *et al.*, *Origins of Human Cancer. Book A. Incidence of Cancer in Humans; Book B. Mechanisms of Carcinogenesis; Book C. Human Risk Assessment*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA, 1977.
4. Kattner P *et al.*, Compare and contrast: Pediatric cancer versus adult malignancies, *Cancer Metastasis Rev* **38**(4):673–682, 2019.
5. Mattiuzzi C, Lippi G, Current cancer epidemiology, *J Epidemiol Glob Health* **9**(4):217, 2019.
6. Wu S, Powers S, Zhu W, Hannun YA, Substantial contribution of extrinsic risk factors to cancer development, *Nature* **529**(7584):43–47, 2016.
7. Follain G, Herrmann D, Harlepp S, Hyenne V, Osmani N, Warren SC, Timpson P, Goetz JG, Fluids and their mechanics in tumour transit: Shaping metastasis, *Nat Rev Cancer* **20**(2):107–124, 2020.
8. Lin M *et al.*, Evolutionary route of nasopharyngeal carcinoma metastasis and its clinical significance, *Nat Commun* **14**(1):610, 2023.
9. Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H, Improving marginal likelihood estimation for Bayesian phylogenetic model selection, *Syst Biol* **60**(2):150–160, 2011.
10. Nave O, A mathematical model for treatment using chemo-immunotherapy, *Heliyon* **8**(4):e09288, 2022.
11. Chroni A, Vu T, Miura S, Kumar S, Delineation of tumor migration paths by using a bayesian biogeographic approach, *Cancers* **11**(12):1880, 2019.
12. Fujii T, Mason J, Chen A, Kuhn P, Woodward WA, Tripathy D, Newton PK, Ueno NT, Prediction of bone metastasis in inflammatory breast cancer using a Markov chain model, *Oncologist* **24**(10):1322–1330, 2019.
13. Lin R-H, Lin C-S, Chuang C-L, Kujabi BK, Chen Y-C, Breast cancer survival analysis model, *Appl Sci* **12**(4):1971, 2022.
14. Zhang B, Probability, Markov chain, and their applications, *J Phys Conf Ser* **1848**:012145, 2021.
15. Ghosh S, Chatterjee ND, Dinda S, Urban ecological security assessment and forecasting using integrated DEMATEL-ANP and CA-Markov models: A case study on Kolkata metropolitan area, India, *Sustain Cities Soc* **68**:102773, 2021.
16. Phelan T, Eslami K, Applications of Markov chain approximation methods to optimal control problems in economics, *J Econ Dyn Control* **143**:104437, 2022.
17. Margarit DH, Romanelli L, A mathematical model of absorbing Markov chains to understand the routes of metastasis, *Biomath* **5**(1):1607281, 2016.
18. World Health Organization International Agency for Research on Cancer, 2023, <https://gco.iarc.fr/today/online-analysis-table>.
19. Welch DR, Hurst DR, Defining the hallmarks of metastasis, *Cancer Res* **79**(12):3011–3027, 2019.
20. Brabletz T, To differentiate or not—routes towards metastasis, *Nat Rev Cancer* **12**(6):425–436, 2012.
21. World Health Organization International Agency for Research on Cancer, 2023, <https://gco.iarc.fr/today/help>.
22. Children’s National, Pediatric Wilms tumor and kidney tumors, 2023, <https://childrensnational.org/visit/conditions-and-treatments/cancer/wilms-tumor-and-kidney-tumors>.
23. Zhang M, Sun J, Bone metastasis from ovarian cancer, Clinical analysis of 26 cases, *Saudi Med J* **34**(12):1270–1273, 2013.
24. Nagasawa M, Johnin K, Hanada E, Yoshida T, Okamoto K, Okada Y, Ueba T, Taga T, Ohta S, Kawauchi A, Advanced childhood testicular yolk sac tumor with bone metastasis: A case report, *Urology* **85**(3):671–673, 2015.

20 Margarit *et al.*

25. Prpić M, Franceschi M, Jukić T, Kust D, Dabelić N, Varjačić T, Lucijanić M, Bolanča A, Kusić Z, Differentiated thyroid cancer in pediatric population: Postoperative treatment with radioactive iodine (i-131), *Acta Clin Croat* **58**(1):119, 2019.
26. Becker M, Stefanelli S, Rougemont A-L, Poletti PA, Merlini L, Non-odontogenic tumors of the facial bones in children and adolescents: Role of multiparametric imaging, *Neuroradiology* **59**(4):327–342, 2017.
27. Vijayasekharan K, Velu U, Godkhindi VM, Munisamy M, *Cancer Therapy, Pediatric Nasopharyngeal Carcinoma*, MedDocs Publishers LLC Online, 2021.
28. González-Motta A, González G, Bermudéz Y, Maldonado MC, Castañeda JM, López D, Cotes-Mestre M, Pediatric nasopharyngeal cancer: Case report and review of the literature, *Cureus* **8**(2):e497, 2016.
29. Irani S, Distant metastasis from oral cancer: A review and molecular biologic aspects, *J Int Soc Prev Community Dent* **6**(4):265, 2016.
30. Jarvis H, Cost NG, Saltzman AF, Testicular tumors in the pediatric patient, *Semin Pediatr Surg* **30**:151079, 2021.
31. Von Allmen D, Malignant lesions of the ovary in childhood, *Semin Pediatr Surg* **14**:100–105, 2005.
32. Heaton TE, Davidoff AM, Surgical treatment of pulmonary metastases in pediatric solid tumors, *Semin Pediatr Surg* **25**:311–317, 2016.
33. Weldon CB, Shamberger RC, Pediatric pulmonary tumors: Primary and metastatic, *Semin Pediatr Surg* **17**:17–29, 2008.
34. Kayton ML *et al.*, Primary lung adenocarcinomas in children and adolescents treated for pediatric malignancies, *J Thorac Oncol* **5**(11):1764–1771, 2010.
35. Ritwik P, Chrisentery-Singleton TE, Oral and dental considerations in pediatric cancers, *Cancer Metastasis Rev* **39**(1):43–53, 2020.
36. Whiteley AE, Price TT, Cantelli G, Sipkins DA, Leukaemia: A model metastatic disease, *Nat Rev Cancer* **21**(7):461–475, 2021.
37. Dörücü D, Şekerci ÇA, Yücel S, Testicular and paratesticular tumors in children, *Bull Urooncol* **20**(4):200–205, 2021.
38. Oktem O, Kim SS, Selek U, Schatmann G, Urman B, Ovarian and uterine functions in female survivors of childhood cancers, *Oncologist* **23**(2):214–224, 2018.
39. Sidney Kimmel Cancer Center-Centers & Clinics-Pediatric Oncology, 2023, https://www.hopkinsmedicine.org/kimmel_cancer_center/cancers_we_treat/pediatric_oncology/becoming_our_patient/cancer_types/lymphoma.html.
40. Allen CE, Kelly KM, Bollard CM, Pediatric lymphomas and histiocytic disorders of childhood, *Pediatr Clin* **62**(1):139–165, 2015.
41. Pediatric Treatment Editorial Board, PDQ Wilms tumor and MD: National Cancer Institute other childhood kidney tumors treatment, Bethesda, 2023, <https://www.cancer.gov/types/kidney/patient/wilms-treatment-pdq>.
42. Sakai S *et al.*, Neuroblastoma with ovarian and pancreatic metastasis, *J Pediatr Surg Case Rep* **73**:101996, 2021.
43. Francis GL *et al.*, Management guidelines for children with thyroid nodules and differentiated thyroid cancer: The American thyroid association guidelines task force on pediatric thyroid cancer, *Thyroid* **25**(7):716–759, 2015.
44. Ordóñez Pereira J *et al.*, Thyroid surgery in pediatric patients: Causes and results, *Cir Pediatr* **34**:9–14, 2021.
45. Stanford Childrens, Brain tumors in children, 2023, <https://www.stanfordchildrens.org/es/topic/default?id=brain-tumors-in-children-90-P02745>.
46. Jhon Hopkins Medicine, Brain tumors in children, 2023, <https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor/pediatric-brain-tumors>.

47. Modica G, Poggiolini L, *A First Course in Probability and Markov Chains*, John Wiley & Sons, 2012.
48. Weiss L, Comments on hematogenous metastatic patterns in humans as revealed by autopsy, *Clin Exp Metastasis* **10**:191–199, 1992.
49. Disibio G, French SW, Metastatic patterns of cancers: Results from a large autopsy study, *Arch Pathol Lab Med* **132**(6):931–939, 2008.