

INTRODUCTION

Properties defining cancerous cells are thought to develop through the effects of somatic mutations accumulated within a cell over multiple generations. The linear behavior of the loglog cancer incidence with age is indicative of a multi-step process where the slope n of the line implies that at least $n + 1$ steps are required to describe the overall incidence.

These genetic alterations are accounted for in this model by considering the mutation status of N genes. Each of these genes can either be normal (0) or altered (1) and the genetic state is defined by the binary sequence of its genes.

A cell is constantly at risk of undergoing a genetic alteration where K_j^i is the mutation rate between the i^{th} and j^{th} states. These rates determine the likelihood that a given cell resides in the i^{th} genetic state with time $P_i(t)$ according to

$$\frac{d\mathbf{P}(t)}{dt} = \begin{bmatrix} -\sum_j k_1^j & k_2^1 & \dots & k_{2^N}^1 \\ k_1^2 & -\sum_j k_2^j & & \vdots \\ \vdots & & \ddots & k_{2^N-1}^{2^N-1} \\ k_1^{2^N} & \dots & k_{2^N-1}^{2^N} & -\sum_j k_{2^N}^j \end{bmatrix} \mathbf{P}(t)$$

In addition, colorectal cancer is unique in that precancerous neoplastic (adenoma) states can be observed and removed. These polyps can share many genetic characteristics with colorectal cancers and the differences between the two gives key insights into the adenoma to carcinoma transformation.