Markovian model of growth and histologic progression in prostate cancer

R. Peirolo[1,*] and M. Scalerandi[2,†]
[1]*Centro Meteorologico Regionale dell'Aereonautica Militare, Linate, 20138, Milano, Italy*
[2]*INFM, Dip. Fisica, Politecnico di Torino, 10124, Torino, Italy*
(Received 4 December 2003; published 2 July 2004)

Models, based on bio-physical and biological considerations, may be very helpful as support tools for traditional diagnostic methodologies and interpretation of statistical data in oncology. This is particularly true when the neoplastic progression and differentiation are rather simple and regular, such as in the case of prostatic adenocarcinomas. Using clinical data as a "statistical ensemble," we propose here a Markovian model to forecast the tumor progression. After validation with clinical data, the model is applied to the determination of the temporal evolution of the risk of metastasis.

DOI: 10.1103/PhysRevE.70.011902 PACS number(s): 87.18.−h, 87.10.+e, 02.70.Rr

I. INTRODUCTION

Prostatic adenocarcinoma, like many other cancers, is characterized in its temporal evolution by volume growth and loss of differentiation [1]. Volume and histologic differentiation (both of a single cell and of the whole tumor) can be quantified in terms of grading systems [2]. Therefore, the pair volume-grade constitutes the most widely used and significant prognostic variable, also because they allow one to predict the presence of metastasis better than any other morphologic or clinical features [3].

Indeed, given the tumor volume and grade, the prediction of its evolution and risk of metastasis is a crucial question, particularly for the case of prostate cancers, which are very frequent in elderly men (more than 30% among men over 50, with a sharp increase with age [4–6]), but very slow in their progression. As a result, there are many more men dying for other reasons, but with a prostate cancer which is often still clinically insignificant, than men that die of prostate cancer [5,6]. Therefore, there is a wide debate about the usefulness of aggressive treatments (surgery, radiation therapy, hormonal therapy), which often cause incontinence and impotence, for tumors that, in most cases, would not threaten the life and health of the patient in his remaining lifetime [4–6]. Even though different studies suggest the strategy of "watchful waiting," giving the patient the appropriate treatment only if and when the tumor is really dangerous [7–10], the decision of optimal scheduling of the screening procedure after detection is still an open question [6].

To help in this kind of decisions mathematical models, based on bio-physical, chemical and biological considerations, may be valuable tools to support and complement traditional diagnostic approaches. Indeed, in the last decades, several models have been proposed to describe the dynamics of tumors [11–16] and angiogenesis [17–19]. Despite the significant conclusions and the validation through comparison with experimental data, two major drawbacks affect most of these models: the large number of parameters (often difficult to evaluate) and the lack of "*in vivo*"/clinical data representing a temporal evolution of the neoplasm.

The case of prostatic adenocarcinoma allows one to partly avoid such difficulties. In fact, as mentioned before, volume and grade are the natural choice for the construction of mathematical models, in which few parameters are required, thanks to the simplicity and regularity of prostate cancers progression. Also, since prostate adenocarcinomas are very diffused and often detected only after the patient's death, a large amount of data about untreated tumors is available. Even though they do not strictly constitute a temporal sequence, they may be used as a statistical "ensemble," which may provide a somewhat equivalent information as a time series.

The most suitable mathematical approach to couple the temporal evolution with the statistical ensemble is to treat growth and histologic progression of the tumor as a "Markovian process," in which the many elements, which influence the dynamics, are accounted for through a stochastic approach [20,21], as discussed in Sec. II. In Sec. III, we estimate the values of the parameters of our model by comparison with autopsy data [1,4]: death can be considered as a particular random sampling of the growth process, and the probabilistic distribution in terms of volume and grade from autopsies can be compared with that generated by a series of simulations. In Sec. IV we consider the risk of metastasis, by introducing the probability of metastasis per year. The probability that a tumor of given mass and grade has already produced metastasis is calculated and compared with experimental data from radical prostatectomies and successive follow-ups [3]. Finally, in Sec. V, we apply our model to the prediction of the temporal evolution and risk of metastasis for a tumor of given initial volume and grade (both affected by uncertainty).

II. THE MODEL

A. Biological considerations

For prostatic adenocarcinomas, the more widely used grading system in clinical applications is that of Gleason [2,6,22,23]. In such a scheme, cancer cells are divided into five categories, starting from the better differentiated (Glea-

*Electronic address: riccardo.peirolo@libero.it
†Electronic address: marco.scalerandi@infm.polito.it

son grade 1) to the worst differentiated (Gleason grade 5). However, in the literature, cells are sometimes grouped in only three categories: well differentiated (Gleason grades 1 and 2), moderately differentiated (Gleason grade 3) and poorly differentiated (Gleason grades 4 and 5) [4,6,8]. The whole tumor is then classified on the basis of the prevailing grade.

In our model we consider only two kinds of cells: well-moderately differentiated (WMD) cells, corresponding to Gleason grades 1–3, and poorly differentiated (PD) cells, Gleason grades 4 and 5. The grade of the tumor is then measured by a continuous variable, given by the percentage of the whole mass that is poorly differentiated. Beside the obvious simplification from a mathematical and computational point of view, our assumption is justified for several reasons. First of all, the available experimental data of volume-grade distributions [1] are too sparse for a comparison with more than two grade classes. Indeed up to now only models with a single kind of cell have been adopted. McNeal [24] and Schmid *et al.* [25] consider an exponential growth for the whole tumor, while Fucks *et al.* [26] consider a Gompertzian one. However, considering a single type of cell seems to be insufficient to grasp the complexity of the growth mechanisms, as shown by Stamey *et al.* [3]. In particular, they demonstrate that the percentage of poorly differentiated cells in tumors from radical prostatectomies is more predictive of the risk of metastasis than the traditional Gleason score.

Furthermore, we distinguish between a Gompertzian and an exponential growth laws for WMD and PD cells, respectively. Such choice is dictated by experimental evidence about the growth rates. Schmid *et al.* [25] suggest an exponential growth with median doubling time of about 70 months for organ confined tumors without any evidence of Gompertzian slowing and a decrease to 43 months for tumors in advanced stages (not organ confined). Nevertheless, they consider only clinically detectable tumors, of at least 0.2 cm^3. The same growth rate cannot apply to microscopic tumors. In fact about 30 doublings are needed to reach a size of 1 cm^3 from a single cell: with a 70 months doubling time, this would be equivalent to about 140 years! Therefore we believe that microscopic tumors formed by well-differentiated cells grow much faster in the initial stages. Eventually, later the growth process slows down because WMD cells are incapable of angiogenesis and only the appearance of PD cells, probably more resistant to apoptosis, determines an increase of the growth constant in more advanced tumors.

B. Mathematical description

To implement our model, we discretize time with a step τ. At each time step, the tumor is described by two variables: the volume v_0 of WMD cells and v_1 of PD cells. We call $V=v_0+v_1$ the total volume. As in [3], we define the grade of differentiation G of the tumor:

$$G = v_1/V. \tag{1}$$

If $G<0.5$, the tumor is considered WMD, otherwise PD.

As mentioned before, WMD cells proliferate following a Gompertzian-like dynamics (see also [27] where a universal Gompertzian-like dynamics is postulated for all tumors in early stages). Hence v_0 grows according to

$$\dot{v}_0 = \frac{v_0(t+\tau) - v_0(t)}{\tau} = \Gamma_0[1 - \exp(-V_c/V)]v_0(t), \tag{2}$$

where $\Gamma_0(>0)$ is the growth rate and V_c is a critical volume at which the slowing of the growth starts becoming evident, due to limitation in nutrient availability [11] and/or cells deformation [28]. The volume of PD cells grows exponentially with growth constant Γ_1:

$$\dot{v}_1 = \frac{v_1(t+\tau) - v_1(t)}{\tau} = \Gamma_1 v_1(t). \tag{3}$$

Finally, the de-differentiation process is described as follows. At each time step, the volume v_0 is divided in N identical volume units (labeled as v_{0i}), each of 0.1 cm^3, with the rest discarded. Each volume unit can transform to v_1 with probability p_t. Hence,

$$\begin{aligned} v_0 &\to v_0 - \sum_{i=1}^{N} v_{0i} r_i, \\ v_1 &\to v_1 + \sum_{i=1}^{N} v_{0i} r_i, \end{aligned} \tag{4}$$

where r_i are random numbers which assume the values 1 (or 0) with probability p_t (or $1-p_t$).

The averaged (deterministic) treatment of the stochastic growth processes, resulting in Eqs. (1) and (2), is justified by the huge number of cells constituting the tumor mass (of the order of 10^9 per cm^3). On the contrary, the slowness of prostate cancer progression suggests that only a small subset of mutations produce de-differentiation, while most of them generate clones not competitive with or with similar growth characteristics of nonmutated cells. (Note that the WMD phase includes a very heterogeneous set of cells, which might have very different properties but follow the same growth law.) As a consequence, even if the time and volume scales used here (about 1 year and 0.01 cm^3) suggest that mutations occur with a high frequency, the de-differentiation process remains a rare phenomenon. As a consequence, the explicit introduction of a random term is not avoidable.

C. Implementation

The model proposed in the previous subsection is implemented as follows:

(1) A population of M different "virtual" patients is considered. At any time t during the evolution, each case i ($i=1,\ldots,M$) is defined by a vector formed by volume, grade and age a of the patient:

$$\vec{w}_{i,t} = \{V_i(t), G_i(t), a_i(t)\}. \tag{5}$$

Of course $a_i(t)=a_i(0)+t$, where $a_i(0)$ is the initial age of the patient.

(2) The initial distribution of patients $\vec{w}_{i,0}$ is determined from autopsy data.

(3) Each case is allowed to evolve up to an age $a_i = T_e$, i.e., for a number of time steps $T_i = [T_e - a_i(0)]/\tau$.

(4) The full set of vectors is then sampled into intervals of width $\Delta V, \Delta G, \Delta a$ to obtain the probability density function to detect a tumor with volume between V and $V+\Delta V$, grade between G and $G+\Delta G$ in a patient with age between a and $a+\Delta a$. Since, for validation, we consider autopsy data, each vector is weighted by a function $h(a_i(t))$, which accounts for the rate of death at age $a_i(t)$.

In the present implementation, the growth parameters $(\Gamma_0, V_c, \Gamma_1, p_t)$ are assumed to be the same for each ensemble member (virtual patient). A better description should of course take into account that the mutation probability which causes de-differentiation depends on genetic predisposition and on the local environment, different from individual to individual. Hence, it might be reasonable to use a distribution of transformation rates (i.e., $p_t = p_{t_i}$) and characterize the transition probability by the resulting average value and variance. Nevertheless, as we will show later, the details of the distribution do not substantially influence the ensemble averages, remaining crucial only for what concerns the particular evolution of each individual member of the ensemble.

III. COMPARISON WITH EXPERIMENTAL DATA

To apply the proposed approach, we discretize time with a one year step, small enough to appreciate some details of the dynamics, which develops over 40-50 years, but long enough to allow us to neglect features over a short time scale (like cellular events). Also, we consider $M = 10\,000$ and $T_e = 100$ years. The latter is large enough so that $h(a(t) > T_e) \sim 0$.

We start our simulations with tumors composed of WMD cells $[G_i(0) = 0]$ and volume $V_i(0) = 10^{-3}$ cm^3 $(\forall i)$. $V_i(0)$ is the lowest detectable size of a tumor in autopsy or from a surgical specimen [5,29]. Considering smaller tumors, although desirable, is not meaningful due to the lack of data for estimating the distribution of tumor occurrence with age.

In order to reproduce as well as possible the experimental distribution for the tumor occurrence age, we consider the distribution of prostate tumors and life tables for the year 2000 in the United States [30]. In Fig. 1, we report the frequency of prostate cancer in every decade of life over 40 derived from a metanalisis of eight series of autopsies [4]. An exponential fit of the form $c(a) = K \exp(a/\alpha)$ with $K = 2.37$ and $\alpha = 28.5$ years, gives the percent of prostate cancer affected people at age a. Although the adopted exponential fit constitutes a somewhat arbitrary extrapolation of cancer occurrence at both young and old ages, we believe that neglecting in the simulations cancer occurrences at $a < 40$ will cause an error larger than the one introduced by a possibly poor extrapolation of the $c(a)$ curve.

The time derivative defines the fraction of men of age a that get a new cancer. If $l(a)$ is the survival probability at age a (taken from [30]), the probability that a prostate tumor starts at age a is (see Table I)

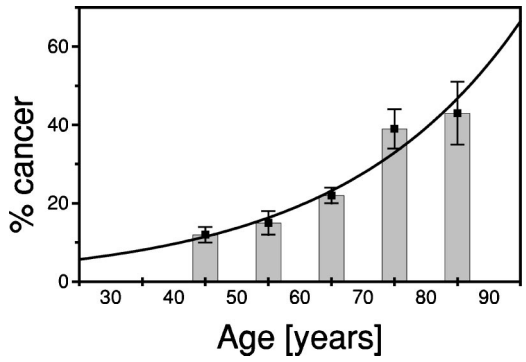

FIG. 1. Frequency of prostate adenocarcinomas in adult males derived from autopsy data [4]. For each decade, we report the percent of cases with a tumor. The solid line represents an exponential fitting.

$$f(a) = \frac{\frac{dc(a)}{da} l(a)}{\sum_a \frac{dc(a)}{da} l(a)}. \tag{6}$$

[Note that the rate of death at the age a is $h(a) = -[dN(a)/da] = -[dl(a)/da]N(a=0)$. In fact, $N(a) = l(a) \cdot N(a=0)$ is the number of living members at age a.]

The initial distribution of ages $a_i(0)$ has been selected from 25 to 99 years, reproducing the distribution described by $f(a)$.

Once the system is let to evolve, the set of $N = \Sigma_i T_i$ vectors yields the distribution (volume, grade and age) of detectable tumors. Integrating over age, we obtain data comparable with the experimental frequencies $(F(V,G)_{\text{exp}})$ obtained from 100 autopsies [1] which are reported in Table II. Experimental data have been rearranged to consider only two grade classes, WMD and PD, as explained in Sec. II.

To obtain the frequencies for the simulated data, we sample the volume into the same five classes as in the statistical analysis of the experimental data. The optimal model parameters have been selected in order to minimize the

TABLE I. Survival probability, new cancers fraction, and cancer occurrence distribution as a function of age. $l(a)$ is taken from [30], $dc(a)/da$ is derived from the exponential fit of Fig. 1 and $f(a)$ from Eq. (6).

Age (a)	Survival probability (%) $l(a)$	New cancers (%/year) $dc(a)/da$	$\mathbf{f(a)}$ (10^{-2})
25	97.7	0.20	0.63
35	96.3	0.28	0.88
45	93.8	0.40	1.22
55	88.7	0.57	1.63
65	77.9	0.81	2.03
75	57.3	1.15	2.13
85	27.3	1.64	1.44
95	4.3	2.32	0.32

TABLE II. Frequencies of detection of a tumor with given grade and volume. Comparison between simulation and experimental results. In parentheses, we report the results obtained by using a patient dependent transition probability with random uniform distribution around p_t.

	WMD		PD	
V (cm^3)	exp	sim	exp	sim
$V<0.05$	19	19.3 (19.2)	1	0.0 (0.0)
$0.05<V<0.17$	18	18.6 (18.6)	2	0.5 (0.5)
$0.17<V<0.46$	18	23.6 (23.9)	2	2.2 (2.4)
$0.46<V<1.40$	13	14.4 (15.3)	7	7.3 (7.4)
$V>1.40$	7	0.2 (0.2)	13	13.9 (13.5)

square distance between the simulated and experimental frequency distributions, considering a Poissonian error for each value of the simulated frequency greater than 5 and an error equal to $\sqrt{5}$ otherwise. We found a minimum for the values of the parameters shown in Table III. Note that $\Gamma_1=0.16$ corresponds to a doubling time of about 56 months, in good agreement with the average of the experimental observations for organ-confined and non-organ-confined tumors (43 and 69 months, respectively) [25].

In Table II we report also the comparison between experimental data and the simulation results. We can observe that the general agreement is satisfactory, confirming that the simple model adopted is sufficient to reproduce enough information about the dynamics of the system. Some discrepancies may be found for big WMD tumors (with $V>1.4$ cm^3) and small PD tumors (with $V<1.7$ cm^3). However, it should be noted the poor statistics (few experimental cases) for such conditions, which indeed correspond to the extreme behaviors of "less harmful tumors that grow in size without histologic progression" and "very aggressive tumors that turn very soon to poorly differentiated patterns." The results reported in parentheses refer to simulations in which the de-differentiation probability is assumed to be patient (marked with the index i) dependent to simulate different environmental conditions for different ensemble members [$p_{t_i}=p_t(0.8+0.2x)$ where x is a random number with uniform distribution between 0 and 1]. As expected, they are very close to the ones obtained with a fixed rate equal for all individuals (results not in brackets).

As already mentioned previously, the extension of the exponential extrapolation of the frequency of prostate cancers to ages below 40 is somewhat arbitrary (see Fig. 1 and relative discussion), but unavoidable, due to the lack of clinical data. To prove that our results are not too sensitive to the choice adopted, we have simulated the ensemble behavior assuming that no cancer starts at ages below 35 [i.e., we impose a cut-off by neglecting in the sampling process the contributions from patients with $a(0)<35$]. We obtain results (see Table IV) in good agreement with the ones reported in Table II, hence justifying our extrapolation. The lower percentages of large volume tumors found when the cut-off is applied are reasonable, considering the shorter (on average) duration of the simulated neoplasm evolution.

TABLE III. Optimal parameters for the simulation as obtained by minimizing the square distance between simulated and experimental distributions as reported in Table II.

Γ_0	V_c	Γ_1	p_t	p_m
1.3	0.018	0.16	0.018	0.022

TABLE IV. Frequencies of detection of a tumor with given grade and volume. Simulation results obtained with a cut-off of cancer occurrence at age 35. In parentheses, we report for reference the simulation results already reported in Table II.

V (cm^3)	WMD	PD
$V<0.05$	19.8 (19.3)	0.0 (0.0)
$0.05<V<0.17$	18.9 (18.6)	0.6 (0.5)
$0.17<V<0.46$	24.1 (23.6)	2.3 (2.2)
$0.46<V<1.40$	14.8 (14.4)	7.2 (7.3)
$V>1.40$	0.1 (0.2)	12.2 (13.9)

To demonstrate the need of a two-populations model, we have also tried to reproduce the experimental data with a single population model. In Table V, we report a comparison between the marginal frequencies $f_m(V)$ of the experiment and of the two-species simulation previously described (ge-sim). Marginal frequencies are defined in terms of volume, i.e., summing the results of Table II for each volume over the grade: $f_m(V)=f_{WMD}(V)+f_{PD}(V)$. We compare the results with that obtained when optimization has been performed considering a single cell species: either PD cells (e-sim), following an exponential growth ($\Gamma_1=1.3$, best choice in terms of square distance), or WMD cells (g-sim), with Gompertzian growth ($V_c=0.5$, $\Gamma_0=5.9$, best choice in terms of square distance). The results show that ge-sim fits much better real data than both e-sim and g-sim (Poissonian error 3.9 against 31.5 and 10.2) and we can conclude that a single species model is not adequate to describe prostate cancer growth.

IV. PREDICTION OF METASTASIS

As a further step, we address the problem of metastasis generation. We assume that the probability of metastasis per

TABLE V. Marginal frequencies of detection of cancer with a given volume. Comparison between experimental data and simulations with two (ge-sim) and one (e-sim, g-sim) species.

cm^3	% exp	% ge-sim	% e-sim	% g-sim
$V<0.05$	20	19.3	35.8	18.9
$0.05<V<0.17$	20	19.1	16.1	12.9
$0.17<V<0.46$	20	25.8	7.6	16.2
$0.46<V<1.40$	20	21.7	14.4	33.3
$V>1.40$	20	14.1	26	18.8

TABLE VI. Probability that a tumor of a given volume has produced metastasis.

cm^3	% met exp	% met sim
$0.5<V<2$	14%	6%
$2<V<6$	39%	37%
$6<V<12$	67%	71%
$V>12$	97%	93%

time step p_{met} is a monotonically increasing function of the volume of PD cells. This choice is in agreement with data reported by Stamey *et al.* [3], where a strong relation between the percentage of PD cells and the presence of metastasis is shown.

The function $p_{met}(v_1)$ is obtained from the total probability law. In fact, the probability that a tumor generates the first metastasis while growing from volume v_1 to v_1+dv_1 is given by

$$p_{met}(v_1 + dv_1) = p_{met}(v_1) + p_{met}(dv_1) - p_{met}(v_1)p_{met}(dv_1). \tag{7}$$

Considering that dv_1 is a small volume, we introduce the linear approximation $p_{met}(dv_1)=p_m dv_1$ to obtain

$$p_{met}(v_1) = 1 - \exp(-p_m v_1). \tag{8}$$

Using Eq. (8), the probability of metastasis is calculated as follows:

(1) At each time step t of a simulation a random number $r \in [0;1]$ is generated.

(2) If $r<p_{met}(v_1(t))$, from that time on there are metastasis.

(3) Summing over many simulations we get the probability that a tumor of total volume V has already produced metastasis.

The ideal data for a comparison come in this case from radical prostatectomies and successive follow-ups. In fact in many cases microscopic, undetectable and clinically insignificant metastasis may be present and only in the following years their presence becomes evident.

Experimental frequencies are given in Table VI for four volume intervals [3]. (Note that volumes are much larger here than in previous cases, since we consider experimental data for tumor patients, i.e., in more advanced stages.) We derived the value of p_m (reported in Table III) that minimize the square distance between the simulation and experimental data (see Table VI). Our results are in good agreement with observations, except for the underestimation in the case $V<2$ cm^3, probably due to neglecting, in our model metastasis due to WMD cells, which are predominant in small tumors. Experiment confirms that even WMD cells (albeit with very low probability) may produce metastasis.

V. APPLICATIONS

A. Results

The values in Table III represent the best choice for the parameters of our model, i.e. those that optimize the growth prediction in terms of volume, differentiation and metastasis generation. Given such values and the vector of the initial values (grading, volume and patient age), the evolution of a single tumor can be followed up and its temporal evolution represented both as a traditional time sequence and as a trajectory in the $(V-G)$ Cartesian plane, i.e., neglecting from now on the age variable. Furthermore, at each point of the trajectory, the probability of metastasis can be estimated using Eq. (8).

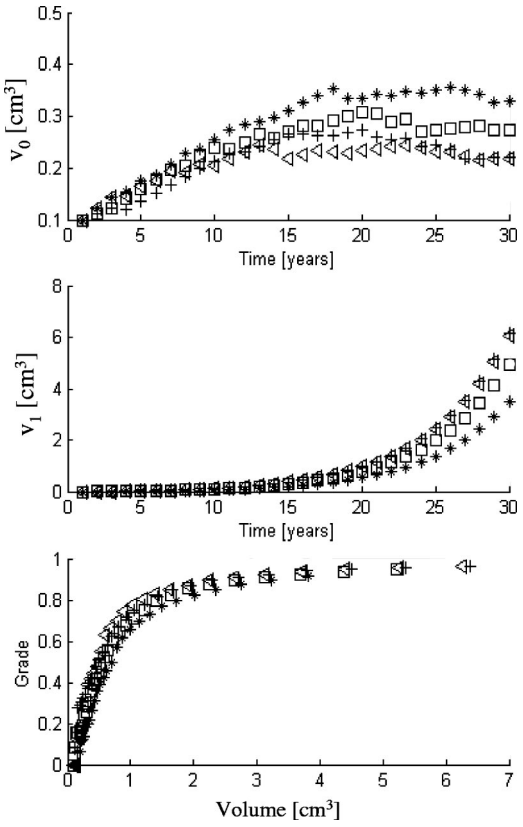

FIG. 2. Temporal evolution of the volumes of WMD and PD cells for four different realizations obtained from the same initial condition and representation of the corresponding trajectories in the volume-grade Cartesian plane.

In Fig. 2, we represent the temporal evolution of four selected cases (from now on called realizations), each starting from the same initial condition: $V=0.1$ cm^3 and $G=0$. As visible from the upper and middle plots (in which the volumes of WMD and PD cells are represented), only slight differences are present. In fact, after an initial phase of proliferation of WMD cells, later de-differentiation becomes dominant and, at long times, well differentiated neoplastic cells reach an asymptotic volume. It is to be noted that, e.g., at 20 years, the total tumor volumes (V) for the four cases are 0.6, 0.8, 1.18 and 1.20 cm^3, which might be a significant difference from a clinical point of view. Also, the cases differ for the time of metastasis occurrence: $t=15$, 12, 18 and 22 years, respectively.

In the lower plot, the same data are represented in a volume-grade Cartesian plot. Here, differences among the realizations are hardly visible.

Data as the ones reported in Fig. 2 are significant only for understanding qualitatively the behavior of the system. A

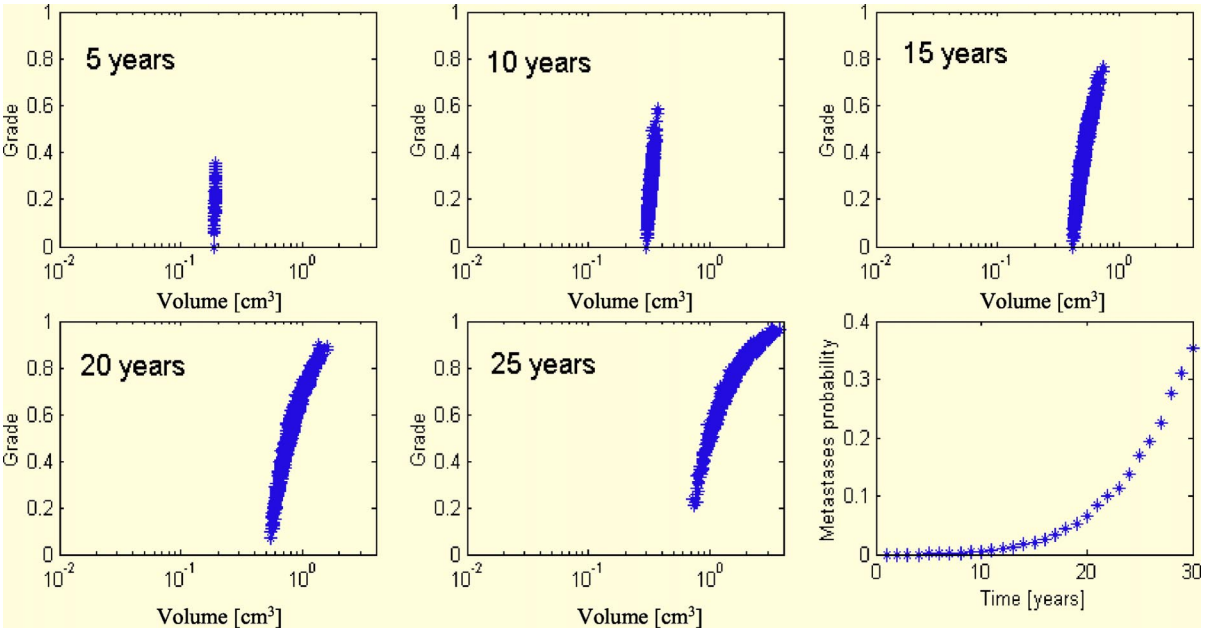

FIG. 3. Forecast of the evolution of a tumor with initial condition: $V=0.1$ cm^3 and $G=0$. Representation of the distribution of volume and grade at 5, 10, 15, 20 and 25 years after initiation (for $M=1000$ realizations) and temporal evolution of the metastasis probability.

different form of analysis, however, allows us to use the model as a diagnostic tool. In fact, for a given starting point $[V(0),G(0)]$, a series of M simulations can be performed, each producing a different path in the $V-G$ plane. At a fixed time t, the M points $(V_i(t),G_i(t))$, where $i(=1,\ldots,M)$ is the trajectory label, form a "cloud" of values, representing a probabilistic forecast of V and G. Also, each temporal evolution may have metastasis from a certain time on. So, summing over all the simulations, we can get the probability of metastasis t years after the initial diagnosis of the tumor.

In Fig. 3, as a typical example, we report snapshots of the distribution in the $V-G$ plane at 5, 10, 15, 20 and 25 years, for $M=1000$ simulations. The initial conditions are the same as the ones used in Fig. 2. As expected, the tumor evolves very slowly. The cloud remains initially quite well localized, denoting the low influence of de-differentiation. Later, 10 years, while the volume remains quite predictable, this is no longer true for the tumor grade. Then, the dispersion for the volume becomes bigger (ranging between 0.6 and almost 2 cm^3 at $t=20$ years). Finally, as soon as the volume turns to be big, generally tumors appear to be poorly differentiated (good predictions for the tumor grade, but low confidence on the tumor volume). It should, however, be noted that the predictability remains good during all stages of the evolution.

As already remarked, such analysis also allows us to evaluate the temporal evolution of the probability of metastasis, as reported in the lower-right plot of Fig. 3. As expected, p_{met} increases exponentially with time, together with the exponential increase of the volume of PD cells (see also Fig. 2).

A more careful analysis, however, requires one to consider that initial conditions are always known with a degree of uncertainty. In fact, except with radical prostatectomy (which of course stops the temporal evolution), tumor volume and differentiation can be guessed only through indirect and imprecise measurements: blood test of prostate-specific antigen (PSA), transrectal ultrasound (TRUS), digital rectal examination (DRE), needle biopsy, etc. [4,6]. So, even the starting conditions are represented not by a single value but by a probabilistic cloud of values $(V_i(0),G_i(0))$, as done in meteorology with the ensemble prediction system (EPS) [31,32], and the temporal evolution causes a further widening of the cloud.

We propose here three virtual experiments starting from the following initial conditions:

(1) T1: a tumor clinically detectable, still small and WMD: $V=0.5$ cm^3, $G=0.25$;

(2) T2: an intermediate tumor: $V=1.5$ cm^3, $G=0.5$;

(3) T3: a big tumor already PD at the diagnosis: $V=4.5$ cm^3, $G=0.75$.

For each of them, we consider a normal distribution of initial values with standard deviation of 0.1 on G and of 0.1 on the logarithm of V. In Fig. 4, we plot the distribution at $t=0$, 5, 10 and 15 years in a $V-G$ plane for the three cases. Except in the case of early diagnosed tumors (small WMD tumors), the probabilistic cloud becomes less spread with increasing time, particularly for what concerns grade, with spreading along the G-axes following a trend comparable to the ones obtained assuming no error in the initial diagnosis (see for comparison Fig. 3). On the contrary, the error in the volume determination seems to affect strongly the predictions. In fact, the indetermination on the logarithm of V remains almost constant when increasing time, with a noticeable indetermination of the tumor volume.

Sensibly different is the case of an early detected tumor (first column). Here, the uncertainty on the initial conditions make the time evolution rather unpredictable. In particular no estimate can be provided about the grade at 5 and 10 years, with obvious consequences on the reliability of the estimate of the metastasis probability.

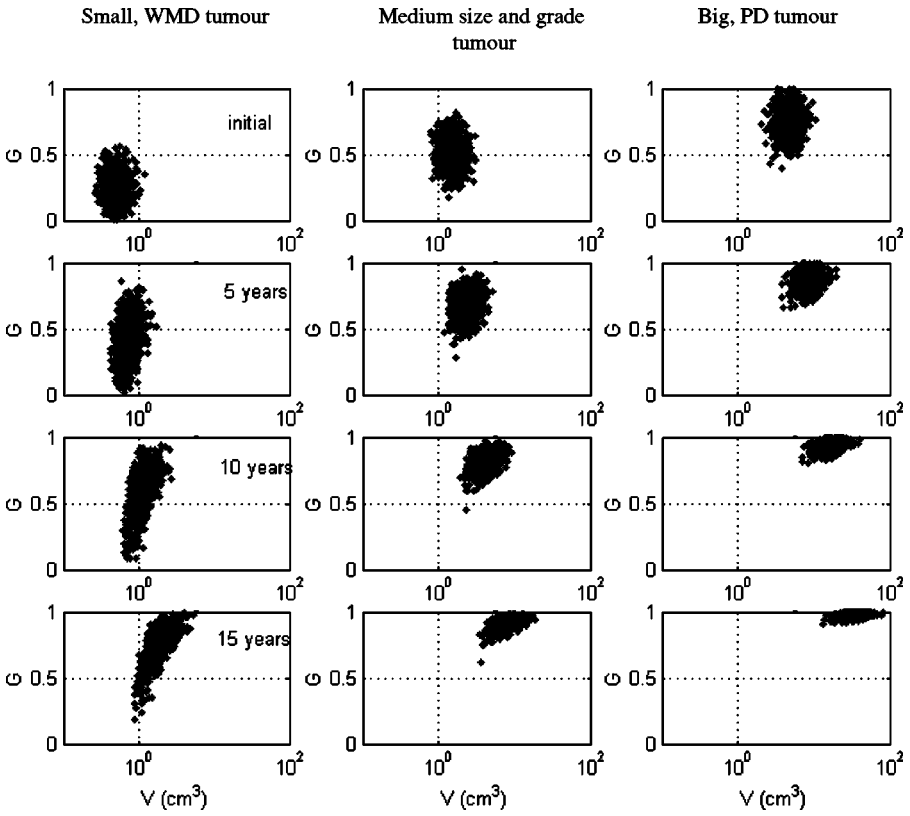

FIG. 4. Effects of the uncertainty on the initial conditions on the forecast of the evolution for three different initial conditions as reported in the text. Representation of the distribution of volume and grade at 5, 10 and 15 years after initiation.

In fact, the same set of data allows us to estimate the probability of metastasis at 5, 10 and 15 years for the three cases considered. Results are reported in Table VII. As we can see, p_{met} is negligible for a T1 tumor at 5 years and still small at 10 years. Even for a T2 tumor p_{met} is low at 5 years and moderate at 10 years, so for the oldest patients surgery or radiation therapy may not be advisable. Our results are in satisfactory agreement with clinical data (reported in parentheses in Table VII), taken from [8,33], even though the correlation between the model initial condition and the experimental diagnosis of the selected cases is rather rough due to the limited information about the experimental data. Also, as already mentioned, our model underestimates slightly the metastasis probability for small WMD tumors, since it does not account for metastatization of WMD cells.

B. Discussion

In 1951 Franks [34] discovered that many elderly men, who had died from different illnesses, had in their prostate a small adenocarcinoma. Since the number of men dying of prostate cancer was much smaller, he concluded that there must be a large group of latent cancers biologically incapable of progression and only a small group of aggressive, harmful prostate cancers. He also proposed that, since it is not possible to distinguish between them, prostate cancer evolution is highly unpredictable. This theory has now been abandoned and there is a wide consensus about a single kind of prostate cancer, whose frequency of inception increases with age and whose growth is very slow in comparison with other tumors [5,25]. As a result, Stamey [35] suggests that prostate cancer is highly predictable. His statement, however, fails to justify the sometimes unexpected outcome of tumor evolution.

TABLE VII. Metastasis probability at 5, 10 and 15 years for the three cases considered in Fig. 4. Data in parentheses refer to clinical data with roughly similar initial conditions.

	5 years	10 years	15 years
T1	3%(7%)	8%(19%)	19%
T2	13%(16%)	35%(42%)	65%
T3	45%(49%)	83%(74%)	98%

Figures 2 and 3 seem to support Stamey's thesis. In fact the tumor dynamics seems to be only slightly stochastic (Fig. 2) and the expected size and grade of the tumor at any given time quite predictable (Fig. 3). Nevertheless, Fig. 4 suggests that the uncertainty in the initial conditions might render predictions less reliable, especially in the case of early diagnosed tumors. In fact, prognosis is often synthesized by a single parameter: the risk of metastasis as a function of time, which can be compared with patient's life expectancy and other clinical data and help in the decision of the appropriate treatment. Such parameter is indeed very dependent on the tumor grade, hence on the quality of the initial conditions.

VI. CONCLUSIONS

In this work, we have proposed a Markovian model to predict volumetric growth and histologic progression in prostate cancers. The model predictions have been validated with a statistical analysis of the relation between initial diagnosis

and final outcome over a group of patients. This approach makes it possible to exploit different sets of data, optimizing the parameters in comparison with them. Here we used autopsy data, considered as a random sampling of the tumor development process, and clinical data about metastasis from radical prostatectomies and follow-up. Temporal evolution of the tumor is represented by a path in the volume-grade plane, the two most significant diagnostic and prognostic variables. Owing to uncertainties both in the initial conditions and in the further evolution, the path is enlarged to a cloud moving with time toward the upper right corner of the plane.

The general agreement of our model with autopsy and clinical data is good, although there are extreme behaviors that a volume-grade model cannot foresee well. Indeed, other factors besides volume and grade, e.g., genetic ones, should be considered for a correct prediction. In particular, aggressive tumors with a high degree of mutation to PD cells and a high probability of metastasis seem to depend on familiar or racial predisposition [1,4] and on environmental conditions. To carefully describe such influence, specifically for an individual patient or for group of patients belonging to a specific class or living in a particular environment, a more accurate analysis of the influence of a random distribution of the parameters (in particular the mutation probability) among members is essential.

We plan to integrate our model with other sets of data such as temporal series of PSA values in untreated patients and information about genetic predisposition. A further step would be to consider not just the total volume but also the growth pattern and directionality (where it starts, whether it penetrates through the capsule, if it reaches the seminal vesicles, etc.). Different tissues have different rigidity and vascularization, leading to different prognosis of further growth and metastasis [3,2]. This can well be reproduced with a local interaction simulation approach (LISA) [36], in which also space is discretized and local mechanisms may be directly included in the model [11,13]. For this kind of description, data from TRUS, DRE and needle biopsy are needed. Since all data are affected by large errors, sophisticated statistical procedures are necessary to produce the best initial conditions, analogous to those that integrate data from ground stations, radiosondes and satellites in meteorology to generate the analysis that initialize models [37].

ACKNOWLEDGMENTS

The authors would like to thank Professor P. P. Delsanto and Dr. M. Griffa (Politecnico di Torino) for useful comments and discussions.

[1] J. E. McNeal, D. G. Bostwick, R. A. Kindrachuk, E. A. Redwine, F. S. Freiha, and T. A. Stamey, Lancet **1**, 60 (1986).
[2] J. E. McNeal, A. A. Villers, E. A. Redwine, F. S. Freiha, and T. A. Stamey, Cancer **66**, 1225 (1990).
[3] T. A. Stamey, J. E. McNeal, C. M. Yemoto, B. M. Sigal, and I. M. Johnstone, J. Am. Med. Assoc. **281**, 1395 (1999).
[4] U. S. Congress, Office of Tecnology Assesment, *Cost and Effectiveness of Prostate Cancer Screening in Elderly Men*, OTA-BP-H-145 (U.S. Government Printing Office, Washington DC, 1995).
[5] T. A. Stamey, F. S. Freiha, J. E. McNeal, E. A. Redwine, A. S. Whittemore, and H. P. Schmid, Cancer **71**, 933 (1993).
[6] M. B. Garnick, Sci. Am. **270** (4), 52 (1994).
[7] J. E. Johansson, L Holmberg, S. Johansson, R. Bergstrom, and H. O. Adami, J. Am. Med. Assoc. **277**, 467 (1997).
[8] G. W. Chodak, R. A Thisted, G. S. Gerber, J. E. Johansson, J. Adolfsson, G. W. Jones, G. D. Chisholm, B. Moskovitz, P. M. Livne, and J. Warner, N. Engl. J. Med. **330**, 242 (1994).
[9] P. C. Albertsen, J. A. Hanley, D. F. Gleason, and M. J. Barry, J. Am. Med. Assoc. **280**, 975 (1998).
[10] L Holmberg, A. Bill-Axelson, F. Helgesen, J. O. Salo, P. Folmerz, M. Haggman, S. O. Andersson, A. Spangberg, C. Busch, S Nordling, J. Palmgren, H. O. Adami, J. E. Johansson, and B. J. Norlen, N. Engl. J. Med. **347**, 781 (2002).
[11] P. P. Delsanto, A. Romano, M. Scalerandi, and G. P. Pescarmona, Phys. Rev. E **62**, 2547 (2000).
[12] D. Drasdo, Phys. Rev. Lett. **84**, 4244 (2000).
[13] B. Capogrosso Sansone, P. P. Delsanto, M. Magnano, and M. Scalerandi, Phys. Rev. E **64**, 021903 (2001).
[14] J. A. Glazier and F. Graner, Phys. Rev. E **47**, 2128 (1993).
[15] A. Bru' *et al.*, Phys. Rev. Lett. **81**, 4008 (1998).
[16] L. M. Sander and T. S. Deisboeck, Phys. Rev. E **66**, 051901 (2002).
[17] B. Capogrosso Sansone, M. Scalerandi, and C. A. Condat, Phys. Rev. Lett. **87**, 128102 (2001).
[18] M. Scalerandi and B. Capogrosso Sansone, Phys. Rev. Lett. **89**, 218101 (2002).
[19] M. J. Holmes and B. D. Sleeman, J. Theor. Biol. **202**, 95 (2000).
[20] L. Chinsan, Stat. Probab. Lett. **40**, 121 (1998).
[21] R. Bartoszynski *et al.*, Math. Biosci. **171**, 113 (2001).
[22] D. F. Gleason, in *Urology Pathology: the Prostate*, edited by M. Tannenbaum (Lea & Febiger, Philadelphia, 1977).
[23] D. F. Gleason, Am. J. Surg. Pathol. **9** (suppl), 53 (1985).
[24] J. E. McNeal, Cancer **23**, 24 (1969).
[25] H. P. Schmid, J. E. McNeal, and T. A. Stamey, Cancer **71**, 2031 (1993).
[26] E. D. Yorke, Z. Fucks, L. Norton, W. Whitmore, and C. C. Ling, Cancer Res. **53**, 2987 (1993).
[27] C. Guiot *et al.*, J. Theor. Biol. **225**, 147 (2003).
[28] M. Griffa, M. Scalerandi, and C. Camagna, Phys. Rev. E (to be published).
[29] A. Wise, T. A. Stamey, J. E. McNeal, and J. E. Clayton, Urology **60**, 264 (2002).
[30] National Vital Statistic Reports 51-3, 9 (2002).
[31] A. J. Simmons *et al.*, Meteorol. Atmos. Phys. **40**, 28 (1989).
[32] F. Molteni *et al.*, Q. J. R. Meteorol. Soc. **122**, 73 (1996).
[33] C. R. Pound, A. W. Partin, M. A. Eisenberger, D. W. Chan, J. D. Pearson, and P. C. Walsh, J. Am. Med. Assoc. **281**, 1591 (1999).

[34] L. M. Franks, J. Pathol. Bacteriol. **68**, 603 (1954).
[35] T. A. Stamey, C. M. Yemoto, J. E. McNeal, B. M. Sigal, and I. M. Johnstone, J. Urol. (Baltimore) **163**, 1155 (2000).
[36] G. Kaniadakis, P. P. Delsanto, and C. A. Condat, Math. Comput. Modell. **17**, 31 (1993); P. P. Delsanto and M. Scalerandi, Phys. Rev. B **68**, 064107 (2003).
[37] R. Daley, *Atmospheric Data Analysis* (Cambridge U. P., Cambridge, 1991).