Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

Estimation of lead-time bias and its impact on the outcome of surveillance for the early diagnosis of hepatocellular carcinoma

Alessandro Cucchetti[1,*], Franco Trevisani[1], Anna Pecorelli[1], Virginia Erroi[1], Fabio Farinati[2], Francesca Ciccarese[3], Gian Lodovico Rapaccini[4], Mariella Di Marco[5], Eugenio Caturelli[6], Edoardo G. Giannini[7], Marco Zoli[1], Franco Borzio[8], Giuseppe Cabibbo[9], Martina Felder[10], Antonio Gasbarrini[11], Rodolfo Sacco[12], Francesco Giuseppe Foschi[13], Gabriele Missale[14], Filomena Morisco[15], Gianluca Svegliati Baroni[16], Roberto Virdone[17], Mauro Bernardi[1], Antonio D. Pinna[1], for the Italian Liver Cancer (ITA.LI.CA) group

[1]*Dipartimento di Scienze Mediche e Chirurgiche, Policlinico S. Orsola-Malpighi, Alma Mater Studiorum – Università of Bologna, Italy;* [2]*Dipartimento di Scienze Chirurgiche e Gastroenterologiche, Unità di Gastroenterologia, Università di Padova, Padova, Italy;* [3]*Divisione di Chirurgia, Policlinico San Marco, Zingonia, Italy;* [4]*Unità di Medicina Interna e Gastroenterologia, Complesso Integrato Columbus, Università Cattolica di Roma, Roma, Italy;* [5]*Divisione di Medicina, Azienda Ospedaliera Bolognini, Seriate, Italy;* [6]*Unità Operativa di Gastroenterologia, Ospedale Belcolle, Viterbo, Italy;* [7]*Dipartimento di Medicina Interna, Unità di Gastroenterologia, Università di Genova, Genova, Italy;* [8]*Dipartimento di Medicina, Unità di Radiologia, Ospedale Fatebenefratelli, Milano, Italy;* [9]*Dipartimento Biomedico di Medicina Interna e Specialistica, Unità di Gastroenterologia, Università di Palermo, Palermo, Italy;* [10]*Ospedale Regionale di Bolzano, Unità di Gastroenterologia, Bolzano, Italy;* [11]*Unità di Medicina Interna e Gastroenterologia, Policlinico Gemelli, Università Cattolica di Roma, Roma, Italy;* [12]*Unità Operativa Gastroenterologia e Malattie del Ricambio, Azienda Ospedaliero-Universitaria Pisana, Pisa, Italy;* [13]*Unità di Medicina Interna, Presidio Ospedaliero di Faenza, Faenza, Italy;* [14]*Unità di Malattie Infettive ed Epatologia, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy;* [15]*Dipartimento di Medicina Clinica e Chirurgia, Unità di Gastroenterologia, Università di Napoli "Federico II", Napoli, Italy;* [16]*Clinica di Gastroenterologia, Università Politecnica delle Marche, Ancona, Italy;* [17]*Dipartimento Biomedico di Medicina Interna e Specialistica, Unità di Medicina Interna, Palermo, Italy*

Background & Aims: Lead-time is the time by which diagnosis is anticipated by screening/surveillance with respect to the symptomatic detection of a disease. Any screening program, including surveillance for hepatocellular carcinoma (HCC), is subject to lead-time bias. Data regarding lead-time for HCC are lacking. Aims of the present study were to calculate lead-time and to assess its impact on the benefit obtainable from the surveillance of cirrhotic patients.

Methods: One-thousand three-hundred and eighty Child–Pugh class A/B patients from the ITA.LI.CA database, in whom HCC was detected during semiannual surveillance (n = 850), annual surveillance (n = 234) or when patients came when symptomatic (n = 296), were selected. Lead-time was estimated by means of appropriate formulas and Monte Carlo simulation, including 1000 patients for each arm.

Results: The 5-year overall survival after HCC diagnosis was 32.7% in semiannually surveilled patients, 25.2% in annually surveilled patients, and 12.2% in symptomatic patients (p <0.001). In a 10-year follow-up perspective, the median lead-time calculated for all surveilled patients was 6.5 months (7.2 for semiannual and 4.1 for annual surveillance). Lead-time bias accounted for most of the surveillance benefit until the third year of follow-up after HCC diagnosis. However, even after lead-time adjustment, semiannual surveillance maintained a survival benefit over symptomatic diagnosis (number of patients needed to screen = 13), as did annual surveillance (18 patients).

Conclusions: Lead-time bias is the main determinant of the short-term benefit provided by surveillance for HCC, but this benefit becomes factual in a long-term perspective, confirming the clinical utility of an anticipated diagnosis of HCC.

Introduction

The global incidence of hepatocellular carcinoma (HCC) is increasing worldwide and the only chance for cure depends on an early diagnosis by means of surveillance of patients at risk [1,2]. The rationale of surveillance is that it can identify HCC at early stages, allowing the use of treatment capable of prolonging survival. In the assessment of the benefit provided by screening or surveillance of any curable disease, lead time represents a potential source of bias [3–5]. Lead time is the time by which

Keywords: Hepatocellular carcinoma; Cirrhosis; Surveillance; Lead-time bias.
Received 24 December 2013; received in revised form 11 March 2014; accepted 24 March 2014; available online 6 April 2014

* Corresponding author. Address: General and Transplant Unit, Department of Medical and Surgical Sciences, S. Orsola-Malpighi Hospital, Alma Mater Studiorum – University of Bologna, Via Massarenti n°9, Bologna, Italy.
E-mail address: aleqko@libero.it (A. Cucchetti).

the diagnosis is anticipated by screening or surveillance with respect to the clinical presentation of a disease [6]. It represents an artificial addition of time to survival of cases detected during screening, leading to a specious improvement in prognosis. Only randomized controlled trials (RCTs) can completely eliminate this bias by comparing mortality rates from the time of patient enrollment in the study, instead of from the time of HCC diagnosis. In the field of surveillance of patients at risk for HCC development, the few available RCTs report conflicting results regarding the benefit of surveillance [7,8], and additional trials are unlikely to be conducted if patients are correctly informed about the risks and benefits of surveillance [9]. Therefore, the actual benefit of surveillance for this type of cancer remains to be defined. Several cohort studies have shown a benefit of surveillance on HCC prognosis, but their results are biased by lead time [10–14]. To limit this bias, some authors have roughly adjusted the survival of patients under surveillance for lead time, but none computed a precise estimation of lead–time bias; therefore, their findings can be substantially affected by different baseline assumptions.

This study aimed at accurately estimating the lead time affecting semiannual and annual surveillance for HCC through a rigorous mathematical model already proposed in other cancer screening programs [15,16]. The impact of lead-time bias on the results achieved by such surveillance programs in a "real world" clinical setting was also explored. Finally, the *number-needed-to screen* (NNS) was calculated to estimate the effect size which should be expected by surveillance.

Patients and methods

Study population

Data were derived from the ITA.LI.CA database. This database currently includes 5136 HCC patients, consecutively seen from January 1987 to December 2012 at 18 medical institutions. Patients having the following inclusion criteria were selected for this study: (a) Child–Pugh class A or B, as surveillance is useless and not recommended by international guidelines in advanced cirrhosis [2,3,12]; (b) treatment description and complete clinical data, and (c) HCC diagnosis reached during surveillance based on liver ultrasonography (US) with or without serum alpha-fetoprotein (AFP) determination, performed every 6 (±1 month; semi-annual surveillance) or 12 months (±1 month; annual surveillance), or at the time of cancer symptom occurrence (outside any surveillance schedule/no-surveillance). Hepatocellular carcinomas incidentally diagnosed as a result of clinical evaluation for other diseases were excluded from the analysis. Patient surveillance was classified as semiannual or annual on the basis of the schedule adopted in the two years preceding HCC diagnosis. In addition, patients under surveillance in whom diagnostic procedures were performed earlier with respect to the scheduled interval, due to the development of signs or symptoms of cancer, were kept in their original surveillance group and computed accordingly. The interval of surveillance was established by the referring physician of each patient. Patients were excluded from the study due to: incomplete clinical data (1195 patients), Child–Pugh class C (278 patients), inconsistent (interval >13 months) surveillance or 3 month-surveillance (total: 808 patients), incidental tumor diagnosis (1475 patients). Accordingly, 1380 patients were enrolled. The time that elapsed between diagnosis and treatment was approximately 40 days for the majority of patients (maximum 2 months) except for candidates for liver transplantation. Cirrhosis was histologically confirmed in 364 patients; in the remaining patients, diagnosis was made unequivocally by clinical and radiological evaluations together with laboratory findings. All patients provided informed consent for the anonymous recording of their data in the ITA.LI.CA database. The study was approved by the Institutional Review Board of each participating center.

Mathematical estimation of lead time

Algebraic details of the mathematical model for lead-time calculation are provided in the Supplementary materials and methods. Briefly, we assumed an *exponential* tumor growth during the *sojourn time* since it best reflects the tumor growth kinetics over the range of sizes at which the majority of HCCs are detected in screening programs (equation 1) [17]. The mean size (together with relevant 95% confidence intervals [95% CI]) of tumors detected during 6-month or 12-month surveillance programs, and the mean size of symptomatic tumors were used for sojourn time calculation (equation 2) [18,19]. Calculation of the sojourn time requires the tumor growth rate to be known, and this variable was derived from the tumor volume doubling time (DT) (equation 3). Thus, the basis of lead-time estimation relies on the doubling time. Hence, a systematic review of the literature was carried out to obtain the most suitable DT values. Details of the literature review are reported in the Supplementary materials and methods. Four studies fulfilled the requirements for the present analysis, involving a total of 155 HCCs, in which the DT was calculated using the formula proposed by Schwartz [20–23]. The distribution of HCC DT was fitted with a log-normal function having $\mu = 4.5253$ and $\sigma = 0.7313$ (Fig. 1). This distribution was used to calculate the *transition rate* to symptomatic disease and lead time, using the appropriate formula (equation 4) [15,16,19].

Simulation methodology

A probabilistic analysis (Monte-Carlo simulation) was initially applied to estimate lead-time and lead-time bias; in this analysis, a theoretical cohort of 1000 patients undergoing semiannual or annual surveillance was considered, and a theoretical cohort of 1000 patients with a symptomatic diagnosis (who did not suffer from lead-time bias) was used as a control group. As previously described, a log-normal distribution was used for doubling time whereas tumor sizes at diagnosis and survival rates varied within a triangular distribution, where interquartile ranges and confidence limits determine the minimum and the maximum values assumed. Base-case time horizon was set at 10 years of follow-up, and a sensitivity analysis was carried out at times varying from 1 year up to 10 years in order to assess the impact of lead time at varying follow-up periods. Survival rates in relationship with surveillance programs were properly calculated and reported in 10-years life-expectancy before and after adjustment for lead-time bias, subtracting the lead time from life-expectancy. Since the whole study population encompasses a large time-period, all the analyses were repeated for patients diagnosed with HCC in more recent years (between 2005 and 2012), on the basis of the premise that advancements in surveillance tools could further anticipate HCC diagnosis [3,4–8,12].

Statistical analysis

Categorical variables are reported in a number of cases and proportions, and comparisons between the subgroups were carried out using the Fisher's exact test. The distribution of continuous variables was checked for normality using the Kolmogorov-Smirnov test, and comparisons between the subgroups were carried out using appropriate tests. Continuous variables are reported as means and 95% CI of the means or as median and interquartile ranges (IQR: 25th and 75th percentiles). Survival rates after HCC diagnosis were computed from the day of diagnosis until death or the last follow-up visit using the Kaplan–Meier method. Survival rates were transformed into monthly probabilities of death, applying the *declining exponential approximation of life expectancy* (DEALE) approach [24].

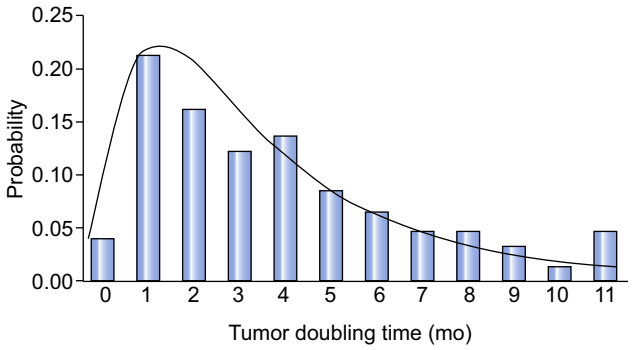

Fig. 1. Distribution of HCC volume doubling time (DT) obtained from the literature review of 155 tumors. The distribution was positively skewed and was fitted with a lognormal function having $\mu = 4.5253$ and $\sigma = 0.7313$ (Median value = 105 days; 25th percentile = 45 days; 75th percentile = 165 days; mode = 45 days). In 53.5% of cases (no = 83), DT value did not exceed 3 months.

Results from the Monte Carlo analyses were reported as NNS. The NNS is calculated as the reciprocal of the absolute risk reduction (NNS = 1/ARR) [25,26] and represents the number of patients who must be enrolled in a screening program over a given time period in order to prevent one death from the given disease. The 5-year overall survival rates were used to compute the NNS. NNS values were compared with previous cost-effectiveness Markov model results, which identified a value of 13 for semiannual US+AFP surveillance and 19 for annual US+AFP surveillance as measurements of the relative effectiveness of the two surveillance strategies [27]. Statistical analyses were carried out with R software version 2.12.0 (R Foundation for Statistical Computing). The Monte-Carlo simulation was carried out using TreeAge-Pro-2008 (TreeAge Software Inc., Williamstown, MA, USA). A two-tailed p value of <0.05 was considered statistically significant in all the analyses.

Results

Of the 1380 cases enrolled, 850 had been diagnosed with HCC during semiannual surveillance (61.6%), 234 during annual surveillance (17.0%), and 296 had a symptomatic diagnosis (21.4%). Thus, the majority (78.4%) of the 1084 patients surveilled was in the semiannual program, and this proportion was adopted in the theoretical cohort of surveilled patients to estimate the lead time of the entire population (see Methods). The median follow-up of the entire study population was 24 months (IQR: 12–46 months). The baseline characteristics of the study population are reported in Table 1. A progressive increase of HCC diagnosed during semiannual surveillance was observed in more recent periods, with a consequent reduction of tumors diagnosed in the annual program or for symptoms. Compared to patients with symptomatic diagnosis, the patients under surveillance more frequently belonged to Child–Pugh class A (p <0.001). They had a greater number of solitary and smaller tumors, and this "stage migration", in turn, increased the proportion of patients undergoing potentially curative treatment, such as transplantation, and resection or ablation (p <0.05 in all cases). A trend toward smaller tumor diameter at diagnosis of semiannual and annual surveillance programs was observed in HCCs diagnosed between 2005 and 2012 with respect to the whole population. As reported in Table 2, the overall survival of patients undergoing surveillance was significantly higher than that of patients with a symptomatic diagnosis. This advantage was confirmed in both Child–Pugh class A and B patients (p <0.05).

Table 1. Baseline characteristics of the study population.

Variable	Semiannual surveillance	Annual surveillance	Symptomatic diagnosis
Number of patients	850 (61.6%)	234 (17.0%)	296 (21.4%)
Period of diagnosis			
1987-1995 (n = 227)	96 (42.3%)	54 (23.8%)	77 (33.9%)
1996-2004 (n = 513)	314 (61.2%)	98 (19.1%)	101 (19.7%)
2005-2012 (n = 640)	440 (68.8%)	82 (12.8%)	118 (18.4%)
Demographic and clinical			
Age (yr)	67.1 (66.5-67.7)†	66.8 (65.6-67.9)†	63.8 (66.6-65.0)
Male sex	595 (70.1%)†	161 (68.8%)†	250 (84.5%)
Hepatitis B	84 (9.9%)*	27 (11.5%)	38 (12.8%)
Hepatitis C	526 (61.9%)†	142 (60.7%)†	102 (34.5%)
Alcohol	80 (9.4%)†	17 (7.3%)†	68 (23.0%)
Multi-etiology	102 (12.0%)*	32 (13.7%)	49 (16.6%)
Other causes	58 (6.8%)†	16 (6.8%)*	39 (13.2%)
Child-Pugh class A	597 (70.2%)†	179 (76.5%)†	155 (52.4%)
Tumor characteristics			
Solitary <2 cm	250 (29.4%)†	37 (15.8%)*	24 (8.1%)
Solitary 2.0-3 cm	198 (23.3%)†	36 (15.3%)†	18 (6.1%)
Solitary 3.1-5 cm	94 (11.1%)	53 (22.6%)†	22 (7.4%)
2-3 nodules <3 cm	147 (17.3%)†	36 (15.4%)*	24 (8.1%)
Outside Milan criteria	161 (18.9%)†	72 (30.8%)†	208 (70.3%)
Size of largest tumor (cm)	2.7 (1.6-3.8)†	3.4 (3.1-3.6)†	4.6 (4.2-5.1)
1987-2004 (n = 740; cm)	2.8 (1.6-3.9)†	3.5 (3.2-3.7)†	4.4 (3.5-4.7)
2005-2012 (n = 640; cm)	2.5 (1.4-3.6)†	3.3 (3.0-3.7)†	4.7 (3.7-5.3)
Treatments			
Transplantation	28 (3.3%)*	5 (2.1%)	2 (0.7%)
Resection	113 (13.3%)	47 (20.1%)	41 (13.9%)
RFA/PEI	370 (43.5%)†	84 (35.9%)†	37 (12.5%)
TACE ± percutaneous ablation	294 (34.6%)	77 (32.9%)	118 (39.9%)
Others/palliation/sorafenib	45 (5.3%)†	21 (9.0%)†	98 (33.1%)

Continuous variables are reported as means and 95% confidence intervals (CI).
p values refer to the pairwise comparison between semiannual or annual surveillance groups to the non-surveillance group: *p <0.05, †p <0.001.
Surveilled patients and patients with a symptomatic diagnosis did not differ regarding presence of any comorbidity, diabetes mellitus or arterial hypertension (data not reported).
RFA, radiofrequency ablation; PEI, percutaneous alcohol injection; TACE, transcatheter arterial chemoembolization.

Table 2. Overall survival rates of the entire study population and stratified by liver function.

Variable	Semiannual surveillance	Annual surveillance	Symptomatic diagnosis
Number of patients	850 (61.6%)	234 (17.0%)	296 (21.4%)
Whole population (n = 1380)			
1-year; %	89.6 (87.2-91.5)†	83.7 (78.2-87.9)†	57.8 (51.2-63.2)
3-year; %	56.8 (53.0-60.4)†	45.9 (39.0-52.5)†	24.2 (20.5-29.4)
5-year; %	32.7 (26.3-39.4)†	25.2 (19.2-32.6)†	12.2 (7.8-16.6)
Child-Pugh A patients (n = 931)			
1-year; %	91.7 (89.0-93.7)†	85.6 (79.4-90.0)†	64.2 (56.0-71.3)
3-year; %	63.3 (58.8-67.5)†	50.1 (42.1-57.6)†	28.5 (21.2-36.1)
5-year; %	36.8 (31.8-41.7)†	28.0 (20.8-35.6)†	16.0 (10.3-22.7)
Child-Pugh B patients (n = 449)			
1-year; %	84.8 (79.6-88.7)†	77.9 (64.4-86.8)*	51.0 (42.4-59.0)
3-year; %	42.4 (35.6-49.0)†	32.8 (20.4-45.7)*	19.6 (13.2-27.0)
5-year; %	24.1 (18.3-30.3)†	16.4 (7.7-27.9)*	7.8 (3.7-14.0)
Patients diagnosed between 2005-2012			
Number of patients	440 (68.8%)	82 (12.8%)	118 (18.4%)
Study population (n = 640)			
1-year; %	92.4 (89.3-94.6)*	83.3 (72.9-89.9)*	64.0 (54.4-72.1)
3-year; %	59.4 (53.3-64.9)*	48.4 (35.5-60.1)*	25.9 (17.3-35.3)
5-year; %	39.0 (31.5-46.5)*	31.4 (18.0-45.6)*	15.3 (8.3-24.3)
Child-Pugh A patients (n = 464)			
1-year; %	94.5 (91.2-96.6)†	83.6 (71.6-90.8)*	66.3 (53.5-76.4)
3-year; %	65.9 (58.9-71.9)†	52.0 (37.5-64.7)*	29.6 (18.0-42.1)
5-year; %	43.2 (33.6-52.4)†	34.8 (18.1-52.2)*	16.2 (7.4-27.9)
Child-Pugh B patients (n = 176)			
1-year; %	85.1 (76.5-90.8)*	82.4 (54.7-93.9)*	58.9 (43.8-71.2)
3-year; %	45.7 (33.9-56.7)*	37.2 (13.1-61.8)	23.3 (11.1-38.0)
5-year; %	26.8 (16.1-38.6)*	18.6 (3.2-44.1)	9.3 (1.0-29.4)

Survival rates are reported together with 95% confidence intervals.
p values refer to pairwise comparison between semiannual or annual surveillance groups to the symptomatic diagnosis group: *p <0.05, †p <0.001.

Lead-time estimation

In a 10-year follow-up perspective, the median lead time calculated for 1000 cirrhotic patients (78.4% of them undergoing semiannual and 21.6% annual surveillance) was 6.5 months (IQR: 3.7–10.3). Distributions of lead time according to surveillance schedules and Child–Pugh classes are reported in Fig. 2. The median lead time of the semiannual program was 7.2 months (IQR: 4.1–12.2) whereas that of the annual program was 4.1 months (IQR: 2.6–6.6). For patients diagnosed with HCC between 2005 and 2012, the median lead time was 7.3 months (IQR: 4.4–12.9) for the semiannual program, and 4.3 months (IQR: 2.7–7.3) for the annual program. The median lead time of the whole study population was 6.5 months for Child–Pugh class A patients (IQR: 3.8–11.1) and 6.4 months for Child–Pugh class B patients (IQR: 3.8–11.0), i.e., unaffected by the functional class.

Effect of lead-time on outcome

The life-expectancy of HCC patients before and after lead-time correction is reported in Table 3. Before correction, both surveillance strategies were associated with an increase in life-expectancy in comparison to a symptomatic HCC diagnosis. Over a 10 year period, the median life-expectancy was 48.5 months for semiannual surveillance, 41.8 months for annual surveillance and 28.5 months for symptomatic diagnosis. The NNS was 8 for semiannual surveillance and 10 for annual surveillance. When stratified by Child–Pugh class, NNS values remained lower (i.e., better) than those indicating the thresholds of effectiveness for the two surveillance strategies [27]. After lead-time correction, the median life expectancy dropped to 41.6 months for semiannual surveillance and to 37.2 months for annual surveillance. The gain in life-expectancy obtainable with both semiannual and annual surveillance was still consistent with an NNS of 13 and 18, respectively. The results changed in Child–Pugh class B patients, in whom lead-time adjustment increased the NNS of both programs above that of the reference values (Table 3). Instead, results did not significantly change in HCC patients diagnosed between 2005 and 2012.

Impact of follow-up length

The results of the sensitivity analysis carried out at varying follow-up lengths are depicted in Fig. 3. After lead-time correction, the differences between life-expectancy associated with surveillance strategies and non-surveillance strategy were clearly dependent on the length of follow-up (Fig. 3A). Namely, surveillance strategies demonstrated an actual benefit over the non-surveillance strategy from the end of the third year of follow-up. In other words, the benefit observed over the first

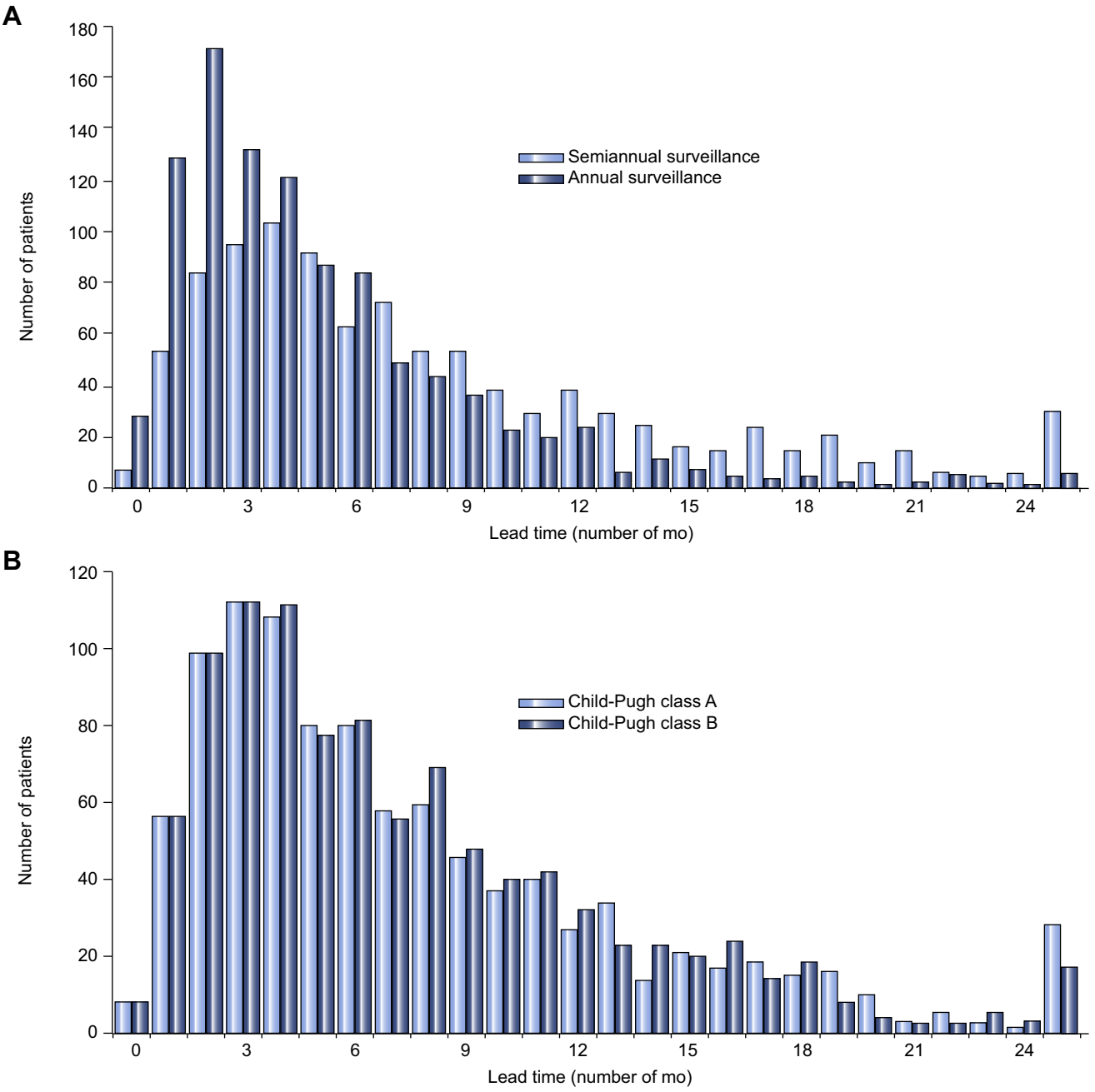

Fig. 2. Lead-time bias estimation. In relation to the surveillance schedule (A) and Child–Pugh class (B).

3 years after HCC diagnosis was apparent, being basically due to lead-time bias. The second sensitivity analysis, at varying follow-up periods, was aimed at identifying tumor DT cut-offs above which surveillance strategies had a negligible benefit (Fig. 3B). As an example of a time point, when considering the third year of follow-up the survival benefit could not be attributable to lead-time bias only for fast-growing HCCs, i.e., those with a DT <120 days in the case of semiannual surveillance and <150 days in the case of the annual schedule. At 5 years of follow-up, the lead-time impact was diluted so that these cut-offs increased to 200 days and 230 days, respectively, indicating that surveillance may also be beneficial for slow-growing tumors. Therefore, very slow-growing tumors give higher lead times and, consequently, need longer follow-ups to overcome the confounding effect of the related bias.

Discussion

In the analysis of results achieved by programs aimed at detecting tumors at early stages, lead-time bias may represent an important confounding factor speciously emphasizing the

Table 3. Results of the Monte Carlo simulation, lead-time correction and number needed to screen calculation.

Variable	Median life-expectancy (months; IQR)	5-year survival	NNS (95% CI)
Before lead-time adjustment			
Whole population			
Semiannual surveillance	48.5 (46.8-50.5)	32.2%	8 (6-10)
Annual surveillance	41.8 (40.1-43.8)	28.4%	10 (8-16)
Symptomatic diagnosis	28.5 (27.2-29.8)	18.2%	[Ref.]
Child-Pugh class A			
Semiannual surveillance	54.5 (53.5-55.7)	36.1%	7 (6-9)
Annual surveillance	45.6 (43.9-47.3)	31.6%	10 (7-15)
Symptomatic diagnosis	31.1 (29.9-32.6)	21.0%	[Ref.]
Child-Pugh class B			
Semiannual surveillance	40.7 (39.6-42.2)	28.6%	7 (6-10)
Annual surveillance	33.9 (31.8-36.2)	23.9%	11 (8-16)
Symptomatic diagnosis	23.1 (22.1-24.3)	14.1%	[Ref.]
After lead-time adjustment			
Whole population			
Semiannual surveillance	41.6 (36.6-44.6)	25.6%	13 (9-26)
Annual surveillance	37.2 (34.3-39.8)	23.8%	18 (11-50)
Symptomatic diagnosis	28.5 (27.2-29.8)	18.2%	[Ref.]
Child-Pugh class A			
Semiannual surveillance	47.5 (42.6-50.6)	31.1%	9 (8-17)
Annual surveillance	40.6 (37.8-43.1)	26.8%	18 (11-49)
Symptomatic diagnosis	31.1 (29.9-32.6)	21.0%	[Ref.]
Child-Pugh class B			
Semiannual surveillance	33.7 (29.1-36.9)	20.6%	15 (11-31)
Annual surveillance	29.0 (25.7-31.9)	18.1%	25 (14-128)
Symptomatic diagnosis	23.1 (22.1-24.3)	14.1%	[Ref.]
Patients diagnosed between 2005-2012			
Child-Pugh class A			
Semiannual surveillance	51.8 (46.0-55.8)	35.5%	8 (6 -11)
Annual surveillance	46.0 (41.1-50.7)	28.7%	15 (10-33)
Symptomatic diagnosis	32.0 (29.9-34.5)	21.8%	[Ref.]
Child-Pugh class B			
Semiannual surveillance	35.2 (29.6-39.3)	21.1%	31 (15-586)
Annual surveillance	32.7 (28.0-38.1)	20.5%	38 (17-134)
Symptomatic diagnosis	27.4 (24.7-30.8)	17.8%	[Ref.]

Median life expectancy is calculated over a 10-year-follow-up.

Numbers needed to screen (NNS) values represent the number of patients who must be enrolled in the surveillance program to prevent one death from the disease in question.

IQR interquartile range (25th and 75th percentiles); CI, confidence interval; NNS, number needed to screen.

observed benefit on survival, if it is not taken into consideration [28,29]. Theoretically, assuming a perfect sensitivity of the surveillance test used, the surveillance interval should be set to correspond to the lead time [18,19]. In this way, almost all HCCs will be detected thanks to the surveillance program. Our study showed that, in the case of surveillance for HCC, lead time varied considerably, within a non-normal distribution, depending on the strategy adopted (Fig. 2). For semiannual surveillance, its median value (7.2 months) was quite similar to the surveillance interval, with the difference attributable to the suboptimal sensitivity of US which increased the lead time [12,14]. Conversely, lead time for annual surveillance (4.1 months) was only one-third of the interval, supporting the opinion that this program is inferior to the semiannual program in terms of tumor stage migration, tumor size at diagnosis and survival [10–13,30]. It should also be pointed out that, as the sensitivity of US increases due to technical advancements, the size of tumors detected by surveillance decreases. Considering this, we repeated the analysis in patients diagnosed with HCC in more recent years, and we found a trend toward a decreased tumor size at diagnosis. This improved diagnostic accuracy of US as a surveillance test results in earlier diagnoses of tumors, but leads to longer, albeit only slightly, lead-time bias. Therefore, the earlier the diagnosis the longer the lead-time bias that has to be expected and this is an important aspect to be taken into account when considering the benefit of surveillance programs.

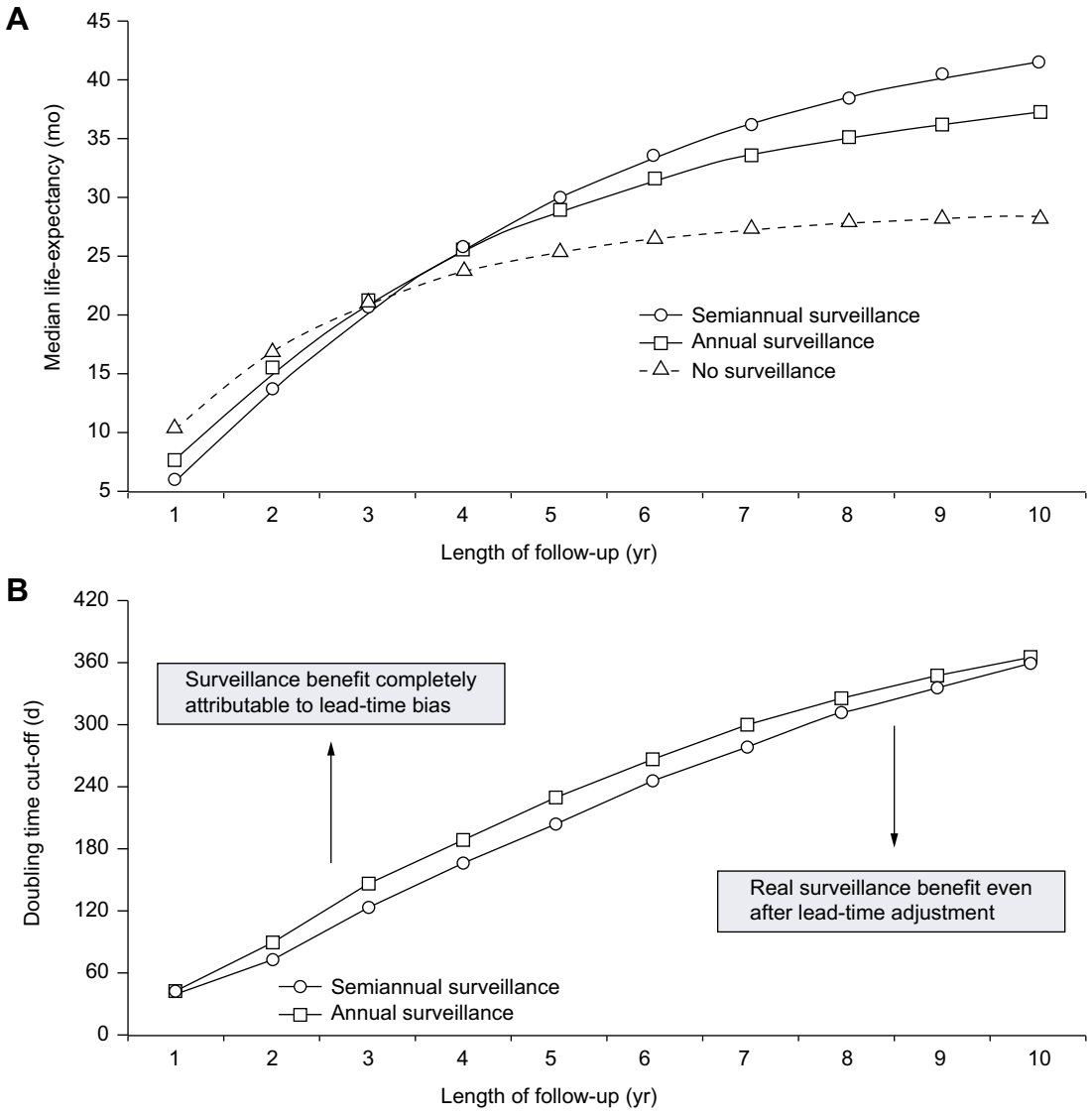

Fig. 3. Sensitivity analyses at varying of follow-up periods. Surveillance determines greater life expectancy, when adjusted for lead-time bias, with respect to non-surveyed patients from the end of the third year of follow-up (A). The second graph (B) weights the impact of the length of follow-up (or survival) and tumor volume doubling time (DT) on the survival benefit of surveillance. Of note, since lead-time bias is strictly dependent on tumor DT (equation 4), this analysis also represents an indirect sensitivity analysis when lead-time varies. If the DT is greater than the thresholds identified for each follow-up time considered, the benefit of semiannual or annual surveillance is fully attributable to lead-time bias. Note that a survival $\geqslant$10 years is required to overcome the confounding effect of the lead-time bias when the DT is approximately 1 year (=very slow-growing tumors). Thus, the longer the DT and the longer the lead-time and the relevant bias that has to be expected.

These deductions, derived from lead-time calculation, are supported by the NNS values obtained in the present study in patients undergoing different surveillance programs for an early diagnosis of HCC (Table 3). The NNS indicates the number of patients to be surveilled to prevent one death (mainly consequent to HCC), and measures the effectiveness of surveillance strategies [25,26]. A previous cost-effectiveness Markov-model analysis identified two values of NNS which define surveillance as effective: 13 for the semiannual program and 19 for the annual program [27]. Any NNS lower than these values indicate that the surveillance strategy is even more effective. In our analysis, semiannual surveillance always resulted in lower NNS values as compared to annual surveillance. Notably, after lead-time correction, which determines a greater decrease in life expectancy for semiannual than for annual surveillance, the more stringent program still maintained lower NNS values. These findings confirmed that 6-month surveillance is more effective than annual surveillance. Nevertheless, this comparison deserves additional comments. First, even annual surveillance produced NNS values lower than the threshold of effectiveness [27] in the entire population and in the subgroup of Child–Pugh A patients, thus proving to be effective in comparison with non-surveillance. Second, in Child–Pugh class B patients, both programs, after lead-time correction, had relatively high NNS values in comparison to previous reports [27] (Table 3), confirming the limited effectiveness of surveillance when it is implemented in patients with advanced cirrhosis [12,31]. These results did not change

when analyzing the subset of patients diagnosed in more recent years.

Some additional findings of our study deserve attention. The survival benefit of surveillance, in relation to lead-time bias, was found to be dependent on the length of follow-up after HCC diagnosis and treatment, and on tumor DT. If the follow-up of a cohort study starting from the time of HCC diagnosis is not long enough, the benefit in life-expectancy observable in surveilled individuals is apparent, being fundamentally attributable to lead-time bias. As a matter of fact, the survival benefit of surveillance becomes factual from the end of the third year of follow-up (Fig. 3A). This information should be considered in order to correctly scrutinize the results of cohort studies dealing with the outcome of surveillance programs for HCC. It can be inferred that only potentially curative treatments, such as resection, ablation, and transplantation (accounting for about 60% of treatments in the individuals surveilled) performed in well compensated patients, can override the confounding impact of lead-time bias, producing reliable results. The second point needing a comment is the direct relationship between lead time and tumor DT (i.e., the longer the DT, the greater the lead time). This is another important point since the distribution of HCC DTs was highly skewed (Fig. 1), and our second sensitivity analysis indicated that the threshold of tumor DT, for considering surveillance as beneficial, is commensurate to the length of follow-up, being rather low if a short follow-up is considered (Fig. 3B) [10]. Over a 3-year period of observation, the survival of patients with HCC diagnosed during surveillance remained significantly better than that of non-surveilled patients if the tumor DT was <3 months (120 days). Conversely, if patient survival is monitored over a long follow-up, such as ten years, lead time will not significantly bias the comparison even in the presence of a tumor DT of about one year. Considering that the review of the literature showed that approximately 90% of HCCs have a DT of less than one year, surveillance programs can lead to a real benefit for the majority of patients at risk of developing this cancer [20–23].

In the present study we tried to estimate, with the greatest possible accuracy, the magnitude of the lead-time bias which can be expected when cirrhotic patients are diagnosed as having HCC during surveillance. Obviously, our mathematical approach was not aimed at replacing RCTs but at providing useful results to correctly interpret the performance of surveillance programs. However, it should be remembered that other biases can affect non-randomized studies on this topic. The most important is length-time bias, which refers to the phenomenon whereby less aggressive, slow-growing tumors have a longer sojourn time and are therefore more likely to be detected by surveillance than faster-growing cancers. Such a selection confers an apparent survival advantage to the surveilled patients, due to the favorable outcome of indolent tumors. A precise estimation of the length-time bias is currently lacking in the setting of early detection of HCC and, hence, dedicated modeling studies are needed to quantify how much this bias artificially improves the performance of surveillance programs.

In conclusion, this study provided a precise estimation of lead-time bias which can be expected in surveillance for HCC according to the surveillance intervals and tumor DT. Considering the impact of tumor DT, an estimation of the surveillance benefit should be assessed by dedicated sensitivity analyses for either slow-growing or fast-growing tumors. Moreover, the confounding role of lead-time bias should be calculated with appropriate formulas capable of handling the length of follow-up, and considering follow-up times long enough to overcome or minimize this bias.

Members of the ITA.LI.CA group

Dipartimento di Scienze Mediche e Chirurgiche, Alma Mater Studiorum – Università di Bologna: Luigi Bolondi, Maurizio Biselli, Paolo Caraceni, Matteo Cescon Marco Domenicali, Francesca Garuti, Annagiulia Gramenzi, Barbara Lenzi, Donatella Magalotti, Fabio Piscaglia, Carla Serra, Matteo Ravaioli, Laura Venerandi; *Unità Operativa di Chirurgia, Policlinico S. Marco, Zingonia*: Paolo Del Poggio, Stefano Olmi; *Unità Operativa di Medicina, Azienda Ospedaliera Bolognini, Seriate*: Claudia Balsamo, Maria Anna Di Nolfo, Elena Vavassori; *Dipartimento di Medicina Clinica e Sperimentale, Università di Padova*: Alfredo Alberti, Luisa Benvegnù, Angelo Gatta; *Dipartimento di Scienze Chirurgiche e Gastroenterologiche, Università di Padova*: Anna Giacomin, Veronica Vanin, Caterina Pozzan, Gemma Maddalo; *Dipartimento di Malattie Apparato Digerente e Medicina Interna, Azienda Ospedaliero-Universitaria di Bologna, Unità Operativa di Radiologia*: Emanuela Giampalma, Alberta Cappelli, Rita Golfieri, Cristina Mosconi, Matteo Renzulli; *Unità Operativa di Gastroenterologia, Ospedale Belcolle, Viterbo*: Serena Dell'Isola, Anna Maria Ialungo, Paola Roselli; *Dipartimento di Medicina Interna, Unità di Gastroenterologia, Università di Genova*: Domenico Risso, Simona Marenco, Giorgio Sammito, Linda Bruzzone; *Unità Operativa di Medicina Interna e Gastroenterologia, Università Cattolica di Roma*: Giulia Bosco, Antonio Grieco, Maurizio Pompili, Emanuele Rinninella, Massimo Siciliano; *Unità di Gastroenterologia, Ospedale Sacro Cuore Don Calabria, Negrar*: Maria Chiaramonte. *Dipartimento di Medicina Clinica e Chirurgia, Unità di Gastroenterologia, Università di Napoli "Federico II", Napoli:* Maria Guarino; *Dipartimento Biomedico di Medicina Interna e Specialistica, Unità di Gastroenterologia, Università di Palermo:* Calogero Cammà, Marcello Maida, Arezia Di Martino, Maria Rosa Barcellona; *Clinica di Gastroenterologia, Università Politecnica delle Marche*, *Ancona:* Laura Schiadà, Stefano Gemini; *Unità di Malattie Infettive ed Epatologia, Azienda Ospedaliero-Universitaria di Parma, Parma*: Elisabetta Biasini, Emanuela Porro; *Unità Operativa Gastroenterologia e Malattie del Ricambio, Azienda Ospedaliero-Universitaria Pisana, Pisa*: Valeria Mismas, Caterina Vivaldi.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Acknowledgments

The authors thank Dott. Alessandro Furlan of the Department of Radiology, University of Palermo, Palermo, for having provided additional data regarding HCC doubling time.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhep.2014.03.037.

References

[1] Altekruse SF, McGlynn KA, Dickie LA, Kleiner DE. Hepatocellular carcinoma confirmation, treatment, and survival in surveillance, epidemiology, and end results registries, 1992–2008. Hepatology 2012;55:476–482.

[2] Giannini EG, Cucchetti A, Erroi V, Garuti F, Odaldi F, Trevisani F. Surveillance for early diagnosis of hepatocellular carcinoma: how best to do it? World J Gastroenterol 2013;19:8808–8821.

[3] European Association for the Study of the Liver. European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2012;56:908–943.

[4] Bruix J, Sherman M. Management of hepatocellular carcinoma: an update – American Association for the Study of Liver Diseases. Hepatology 2011;53:1020–1022.

[5] Omata M, Lesmana LA, Tateishi R, Chen PJ, Lin SM, Yoshida H, et al. Asian Pacific Association for the Study of the Liver consensus recommendations on hepatocellular carcinoma. Hepatol Int 2010;4:439–474.

[6] Walter SD, Stitt LW. Evaluating the survival of cancer cases detected by screening. Stat Med 1987;6:885–900.

[7] Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol 2004;130:417–422.

[8] Chen JG, Parkin DM, Chen QG, Lu JH, Shen QJ, Zhang BC, et al. Screening for liver cancer: results of a randomised controlled trial in Qidong, China. J Med Screen 2003;10:204–209.

[9] Poustchi H, Farrell GC, Strasser SI, Lee AU, McCaughan GW, George J. Feasibility of conducting a randomized control trial for liver cancer screening: is a randomized controlled trial for liver cancer screening feasible or still needed? Hepatology 2011;54:1998–2004.

[10] Wong GL, Wong VW, Tan GM, Ip KI, Lai WK, Li YW, et al. Surveillance programme for hepatocellular carcinoma improves the survival of patients with chronic viral hepatitis. Liver Int 2008;28:79–87.

[11] Han KH, Kim DY, Park JY, Ahn SH, Kim J, Kim SU, et al. Survival of hepatocellular carcinoma patients may be improved in surveillance interval not more than 6 months compared with more than 6 months: a 15-year prospective study. J Clin Gastroenterol 2013;47:538–544.

[12] Trevisani F, Santi V, Gramenzi A, Di Nolfo MA, Del Poggio P, Benvegnù L, et al. Surveillance for early diagnosis of hepatocellular carcinoma: is it effective in intermediate/advanced cirrhosis? Am J Gastroenterol 2007; 102:2448–2457.

[13] Yuen MF, Cheng CC, Lauder IJ, Lam SK, Ooi CG, Lai CL. Early detection of hepatocellular carcinoma increases the chance of treatment: Hong Kong experience. Hepatology 2000;31:330–335.

[14] El-Serag HB, Kramer JR, Chen GJ, Duan Z, Richardson PA, Davila JA. Effectiveness of AFP and ultrasound tests on hepatocellular carcinoma mortality in HCV-infected patients in the USA. Gut 2011;60:992–997.

[15] Duffy SW, Nagtegaal ID, Wallis M, Cafferty FH, Houssami N, Warwick J, et al. Correcting for lead-time and length bias in estimating the effect of screen detection on cancer survival. Am J Epidemiol 2008;168:98–104.

[16] Prevost TC, Launoy G, Duffy SW, Chen HH. Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. Am J Epidemiol 1998;148:609–619.

[17] Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures for mass screening programs. Biometrics 1984;40:1–14.

[18] Duffy SW. Screening, sojourn time. encyclopedia of biostatistics. John Wiley & Sons, Ltd.; 2005. http://dx.doi.org/10.1002/0470011815.b2a04052.

[19] Michaelson J, Satija S, Moore R, Weber G, Halpern E, Garland A, et al. Estimates of the sizes at which breast cancers become detectable on mammographic and clinical grounds. J Womens Imag 2003;5:11–19.

[20] Furlan A, Marin D, Agnello F, Di Martino M, Di Marco V, Lagalla R, et al. Hepatocellular carcinoma presenting at contrast-enhanced multi-detector-row computed tomography or gadolinium-enhanced magnetic resonance imaging as a small (⩽2 cm), indeterminate nodule: growth rate and optimal interval time for imaging follow-up. J Comput Assist Tomogr 2012;36:20–25.

[21] Cucchetti A, Vivarelli M, Piscaglia F, Nardo B, Montalti R, Grazi GL, et al. Tumor doubling time predicts recurrence after surgery and describes the histological pattern of hepatocellular carcinoma on cirrhosis. J Hepatol 2005;43:310–316.

[22] Taouli B, Goh JS, Lu Y, Qayyum A, Yeh BM, Merriman RB, et al. Growth rate of hepatocellular carcinoma: evaluation with serial computed tomography or magnetic resonance imaging. J Comput Assist Tomogr 2005;29:425–429.

[23] Kubota K, Ina H, Okada Y, Irie T. Growth rate of primary single hepatocellular carcinoma: determining optimal screening interval with contrast enhanced computed tomography. Dig Dis Sci 2003;48:581–586.

[24] Beck JR, Pauker SG, Gottlieb JE, Klein K, Kassirer JP. A convenient approximation of life expectancy (the "DEALE") II. Use in medical decision-making. Am J Med 1982;73:889–897.

[25] Rembold CM. Number needed to screen: development of a statistic for disease screening. BMJ 1998;317:307–312.

[26] Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ 1995;310:452–454.

[27] Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Jackson S, et al. Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. Br J Cancer 2008;98:1166–1175.

[28] Zelen M, Fainleib M. On the theory of screening for chronic disease. Biometrika 1969;56:601–614.

[29] Hutchison GB, Shapiro S. Lead time gained by diagnostic screening for breast cancer. J Natl Cancer Inst 1968;41:665–681.

[30] Santi V, Buccione D, Di Micoli A, Fatti G, Frigerio M, Farinati F, et al. The changing scenario of hepatocellular carcinoma over the last two decades in Italy. J Hepatol 2012;56:397–405.

[31] Cucchetti A, Trevisani F, Cescon M, Ercolani G, Farinati F, Del Poggio P, et al. Cost-effectiveness of semi-annual surveillance for hepatocellular carcinoma in cirrhotic patients of the Italian Liver Cancer population. J Hepatol 2012;56:1089–1096.