Int. J. Radiation Oncology Biol. Phys., Vol. 81, No. 2, pp. 537–544, 2011

0360-3016/$ - see front matter

doi:10.1016/j.ijrobp.2011.03.012

BIOLOGY CONTRIBUTION

DEVELOPMENT OF A MULTICOMPONENT PREDICTION MODEL FOR ACUTE ESOPHAGITIS IN LUNG CANCER PATIENTS RECEIVING CHEMORADIOTHERAPY

Kim De Ruyck, Ph.D.,* Nick Sabbe, M.Sc.,† Cary Oberije, Ph.D.,‡ Katrien Vandecasteele, M.D.,§ Olivier Thas, Ph.D.,† Dirk De Ruysscher, M.D., Ph.D.,‡ Phillipe Lambin, M.D., Ph.D.,‡ Jan Van Meerbeeck, M.D., Ph.D.,‖ Wilfried De Neve, M.D., Ph.D.,§ and Hubert Thierens, Ph.D.*

Departments of *Basic Medical Sciences and †Applied Mathematics, Biometrics and Process Control, Ghent University, Ghent, Belgium; ‡Department of Radiation Oncology (MAASTRO Clinic), Research Institute of Growth and Development, Maastricht University Medical Center, Maastricht, The Netherlands; and Departments of §Radiation Oncology and ‖Respiratory Medicine, Ghent University Hospital, Ghent, Belgium

Purpose: To construct a model for the prediction of acute esophagitis in lung cancer patients receiving chemoradiotherapy by combining clinical data, treatment parameters, and genotyping profile.

Patients and Methods: Data were available for 273 lung cancer patients treated with curative chemoradiotherapy. Clinical data included gender, age, World Health Organization performance score, nicotine use, diabetes, chronic disease, tumor type, tumor stage, lymph node stage, tumor location, and medical center. Treatment parameters included chemotherapy, surgery, radiotherapy technique, tumor dose, mean fractionation size, mean and maximal esophageal dose, and overall treatment time. A total of 332 genetic polymorphisms were considered in 112 candidate genes. The predicting model was achieved by lasso logistic regression for predictor selection, followed by classic logistic regression for unbiased estimation of the coefficients. Performance of the model was expressed as the area under the curve of the receiver operating characteristic and as the false-negative rate in the optimal point on the receiver operating characteristic curve.

Results: A total of 110 patients (40%) developed acute esophagitis Grade ≥2 (Common Terminology Criteria for Adverse Events v3.0). The final model contained chemotherapy treatment, lymph node stage, mean esophageal dose, gender, overall treatment time, radiotherapy technique, rs2302535 (*EGFR*), rs16930129 (*ENG*), rs1131877 (*TRAF3*), and rs2230528 (*ITGB2*). The area under the curve was 0.87, and the false-negative rate was 16%.

Conclusion: Prediction of acute esophagitis can be improved by combining clinical, treatment, and genetic factors. A multicomponent prediction model for acute esophagitis with a sensitivity of 84% was constructed with two clinical parameters, four treatment parameters, and four genetic polymorphisms. © 2011 Elsevier Inc.

Prediction, Esophagitis, Radiotherapy, Genetic polymorphisms, Lasso logistic regression.

INTRODUCTION

Lung cancer has the highest incidence and mortality rate of all cancers in Western countries, with most patients presenting with advanced-stage disease at the time of diagnosis (1). The standard of care for locally advanced non–small-cell lung cancer is concurrent chemoradiotherapy. Treatment success is, however, still constrained by poor local control and posttherapy toxicity as acute esophagitis (2, 3).

Numerous studies have attempted to define clinical and dosimetric predictors of radiation-induced esophagitis (4–21). Factors found to correlate with important acute esophagitis include concurrent chemoradiotherapy, lymphatic status, and a number of dose–volumetric parameters. Because the results varied considerably across different institutions, their clinical usefulness remains restricted. Simultaneously, significant research efforts have been made to link genetic polymorphisms in selected genes to radiation-induced toxicity (22–30). The majority of these radiogenetics studies considered patients with breast or prostate cancer, whereas only a few research groups reported on radiation-induced toxicity after radiotherapy for lung cancer (28–31). Because of the limited genetic studies performed for esophagitis and the inconsistent outcomes for other radiation-induced toxicities, genetic biomarkers are not yet implementable in the clinic.

Reprint requests to: Kim De Ruyck, Ph.D., Ghent University, Department of Basic Medical Sciences, Proeftuinstraat 86, B-9000 Ghent, Belgium. Tel: (+32)-926-46656; Fax: (+32) 926-46696; E-mail: kim.deruyck@UGent.be

Supported by Grant No. 365D6706 from the Stichting tegen Kanker and Grant No. 09/PDO/061 from the Fonds voor Wetenschappelijk Onderzoek.

Conflicts of interest: none.

Acknowledgments—The authors thank all the study participants and data managers, Dr. Bart Loeys, and Virginie de Gelder.

Received Jan 25, 2011, Accepted for publication March 8, 2011.

Over the years it has become clear that the effects of dose distribution, clinical parameters, and individual genetic variation on radiation-induced toxicity may not be evaluated separately. As a result, studies are emerging that correct for dosimetric and patient-related risk factors when trying to link genetic polymorphisms with a clinical phenotype after therapy (29, 31–34). However, a model to predict susceptibility to radiation in individual patients is still unavailable. Therefore, we aimed at constructing a predictive algorithm for acute esophagitis by combining clinical data, treatment parameters, and genotypic information. This would enable us to individualize patient treatment.

PATIENTS AND METHODS

Study population

Two hundred eighty-nine lung cancer patients treated with curative radiotherapy between February 2004 and August 2009 were enrolled. Of these, 273 were suited to perform the study (Fig. 1). A total of 213 patients were recruited from the MAASTRO Clinic and 60 patients from the Ghent University Hospital. Clinical data and treatment details are presented in Table 1. The majority of patients were treated with three-dimensional conformal radiotherapy (3D-CRT) as opposed to intensity-modulated radiotherapy (IMRT). The median tumor dose was 60 Gy at 1.5–2.69 Gy per fraction (2 patients received 7.5 Gy per fraction). The details of the different radiotherapy treatment regimens can be found in Appendix E1 (available online). For the MAASTRO Clinic patients, the dosimetric parameters were calculated using a commercial radiotherapy treatment planning system (XiO; Computerized Medical Systems, St. Louis, MO). The dosimetric parameters for the Ghent patients were calculated using an in-house-developed planning system, with a final dose computation using a commercial radiotherapy planning system (Pinnacle; Philips Medical Systems, Best, The Netherlands). In both centers, the esophagus was delineated using the external esophageal contour from the cricoid cartilage to the gastroesophageal junction. Esophageal toxicity was scored using the Common Terminology Criteria for Adverse Events scale version 3.0 (35). Acute esophagitis was defined as dysphagia Grade ≥2 at any time during or at maximum 3 months after radiotherapy treatment. Genomic DNA was obtained from fresh blood (Ghent samples) or frozen buffy coat (MAASTRO samples) using the Puregene genomic DNA purification kit (Gentra Systems, Minneapolis, MN). The study was approved by the ethics committees of both centers, and all study participants provided written informed consent. The MAASTRO Clinic cohort study was filed at clinicaltrials.gov (no. NCT01084785).

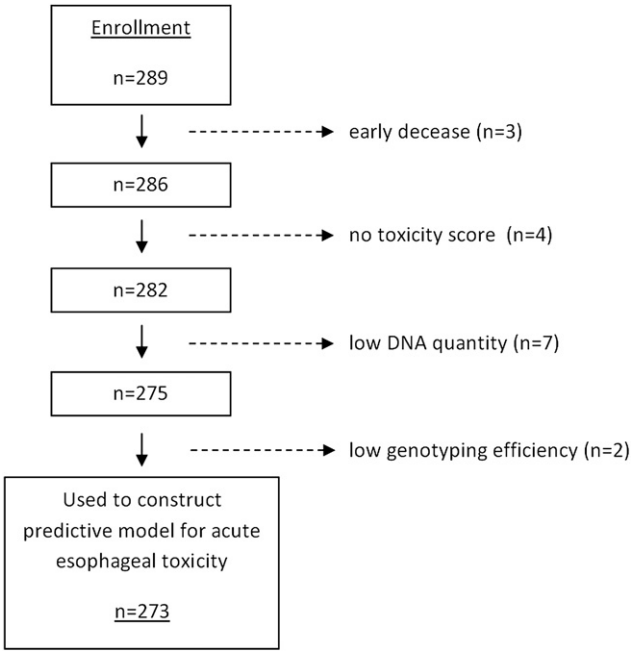

Fig. Study flowchart.

Selection of candidate genes and genetic variations

On the basis of a literature search, a total of 112 candidate genes belonging to the following categories were considered: early response genes ($n = 15$), cytokines and growth factors ($n = 14$), signal transduction of cytokines and growth factors ($n = 53$), adhesion molecules ($n = 9$), extracellular matrix genes ($n = 13$), and others ($n = 8$). Single nucleotide polymorphisms (SNPs) were selected after functional tagging based on evolutionary conserved regions (ECRs) using ECR Browser (36). The 5′ flanking region of each gene (5 kb) was included to thoroughly examine any possible regulatory regions. Conserved regions with a minimal length of 100 bp with minimal 80% equality over four species (mouse, rat, rhesus monkey, and human) were considered. The criteria for picking SNPs in the ECRs were as follows: minor allele frequencies in populations of Caucasian ethnicity >5%, Illumina SNP score >0.60 (proprietary score used to determine the overall success of the assay), and no linkage ($r^2 > 0.80$) with other SNPs. Less-stringent criteria were used for cytokines, their receptors, and cell adhesion molecules (ECRs of 75% equality over three species). The selection was expanded with a number of SNPs with a proven biological function from the literature. Finally, 384 SNPs were retained (Table E1).

Genotyping and quality control

The DNA quality and quantity were checked before genotyping, and 7 individuals were omitted from genotyping because of low DNA quantity. Genotyping was performed using the Illumina Goldengate technology (DNA Vision, Charleroi, Belgium). Upon completion of genotyping, quality control (QC) processes were run to guarantee the accuracy of the genotyped dataset. Two individuals were dropped on the basis of low genotyping efficiency (<90%). Single nucleotide polymorphisms were eliminated on the basis of low reliability of cluster separation, low signal intensity, low call frequency (<75%), absence of a minor allele in the dataset, and deviation from Hardy-Weinberg equilibrium ($p < 0.0001$). After application of stringent QC, approximately 3% of the samples and approximately 13% of the SNPs were eliminated. Overall, QC was completed yielding 273 individuals and 332 SNPs (Table E1).

Statistics

The predicting model was achieved in two steps. First, lasso logistic regression was applied to the full dataset, for which the lasso parameter was chosen so as to maximize the area under the curve (AUC) of the receiver operator characteristic (ROC) curve. The latter was estimated from 10-fold cross-validation (37). Second, the implied set of predictors was passed to classic logistic regression for unbiased estimation of the coefficients.

Lasso regression shrinks the coefficient estimates toward zero, with the degree of shrinkage depending on an additional parameter, lambda (λ). In this way, coefficient estimates can be forced to be exactly zero, thereby effectively eliminating a number of variables.