Table 1. Characteristics of patients by *EGFR* mutation status

	EGFR mutations		*P*
	Present (n = 51)	Absent (n = 103)	
Age (mean ± SD, years)	66.5 ± 10.9	67.8 ± 9.6	ns‡
Gender			
Male	13 (14%)	78 (86%)	$P < 0.001$
Female	38 (60%)	25 (40%)	
Smoking status			
Smoker	12 (13%)	78 (87%)	$P < 0.001$
Nonsmoker	39 (61%)	25 (39%)	
Histology			
Adenocarcinoma	49 (41%)	70 (59%)	$P = 0.02$
Squamous cell carcinoma	1 (4%)	24 (96%)	
Others	1 (10%)	9 (90%)	
Pathological stage†			
I	35 (38%)	58 (62%)	ns
II	3 (21%)	11 (79%)	
III	8 (33%)	16 (67%)	
IV	5 (23%)	17 (77%)	

†Examination for distant metastasis was not done in one case.
‡Not significant.
EGFR, epidermal growth factor receptor.

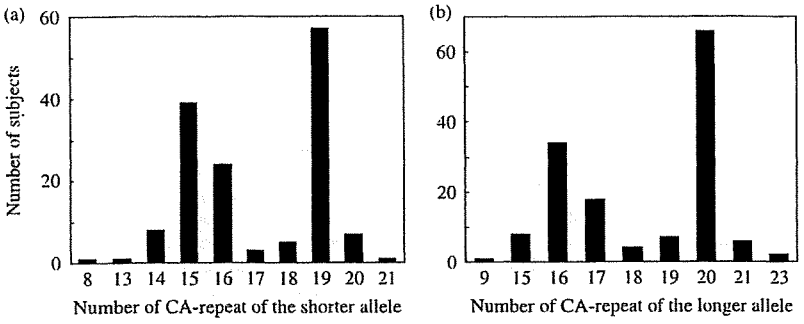

Fig. 1. Distribution of CA-repeat of the shorter allele (a), and the longer allele (b) in lung cancer patients.

All *P*-values presented are two-tailed; those below 0.05 were considered statistically significant. Analyzes were done using SPSS version 12 (SPSS Inc.).

Results

***EGFR* mutations and numbers of CA-repeats.** Clinicopathological characteristics of the 154 patients are shown in Table 1. *EGFR* mutations occurring at exons 18, 19, or 21 were detected in 51 (33%) of the patients. Clear associations of *EGFR* mutations were observed with gender ($P < 0.001$), smoking status (smoker vs non-smoker; $P < 0.001$), and histological type (squamous cell carcinoma, adenocarcinoma, or other; $P = 0.02$). *EGFR* mutations were found in 49 (41%) of 119 adenocarcinoma patients but in only one (4%) of 25 squamous cell carcinoma patients. Distributions of CA-repeat numbers at intron 1 in the shorter and longer alleles of the *EGFR* gene among 146 patients (eight patients were non-informative) revealed two peaks (Fig. 1): CA 19 (39%) and CA 15 (27%) in the shorter alleles and CA 20 (45%) and CA 16 (23%) in the longer alleles.

***EGFR* mutations and CA-repeats.** Because most (96%) of the *EGFR* mutations were found in adenocarcinoma patients, we investigated the relationship between *EGFR* mutations and CA-repeat length only in those patients; 113 patients were examined (we were unable to determine both CA-repeat number and *EGFR* mutation status in six patients). *EGFR* mutations at exons 18, 19, and 21 were detected in 3 (3%), 22 (20%), and 23 (20%), respectively, of the adenocarcinoma patients. The mutations at exon 21 were L858R in 21 patients and L861Q in one patient. All mutations at exon 19 were in-frame deletions with sizes ranging from 9 to 24 base pair.

When study patients were divided into two groups according to the median CA-repeat length (<17 or ≥17) in the shorter allele, *EGFR* mutations at exon 19 were more frequently found in the patients with shorter CA-repeat length ($P = 0.02$, Table 2). On the other hand, *EGFR* mutation frequencies at exons 18 and 21 did not significantly differ with CA-repeat length (Table 2). Logistic regression analysis confirmed the significant contribution of CA-repeat to *EGFR* mutations at exon 19 among the lung adenocarcinoma patients ($P = 0.02$), independently of gender, age, and smoking status (data not shown). In contrast, the same analysis for overall (at exons 18, 19, and 21 combined) *EGFR* mutations in all 154 lung cancer patients demonstrated statistically significant associations with gender and smoking status, but not with CA-repeat length, and the analysis for overall *EGFR* mutation frequency only in lung adenocarcinoma patients failed to demonstrate a statistically significant association with either gender or smoking status (data not shown). An analysis using CA-repeat length in the longer allele produced results similar to, but not as clear as, those obtained for the shorter allele (data not shown); hereafter, only the results for CA-repeats in the shorter allele are shown.

***EGFR* mRNA expression and CA-repeats in non-cancerous lung tissues.** *EGFR* mRNA levels in non-cancer tissues from 74 lung adenocarcinoma patients (excluding 39 non-informative cases),

doi: 10.1111/j.1349-7006.2008.00804.x

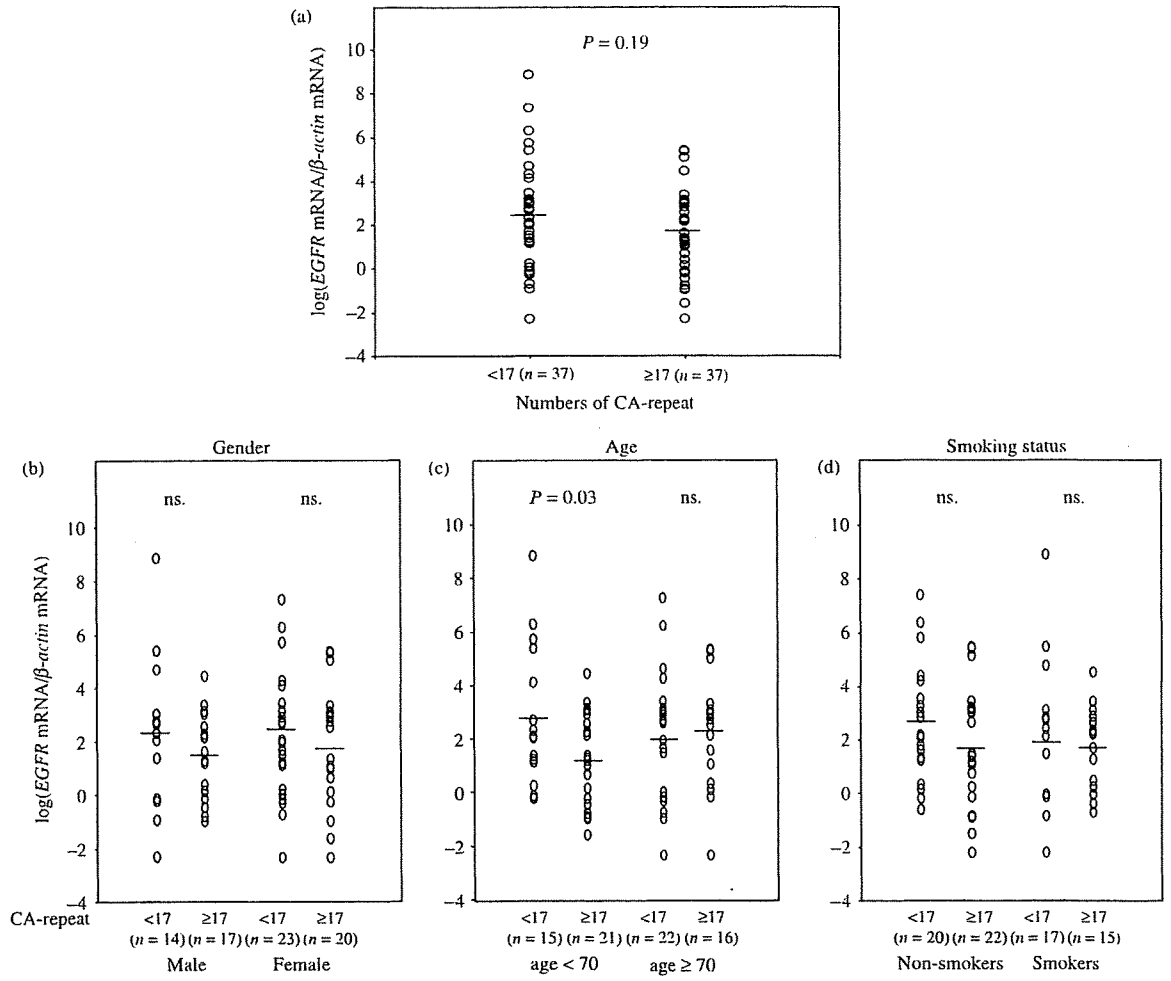

Fig. 2. Distribution of log(*epidermal growth factor receptor (EGFR)* mRNA/*β-actin* mRNA) in non-cancerous tissues among lung adenocarcinoma patients according to number of CA-repeats (<17 or ≥17) (a), gender (b), age (c), and smoking status (d). ns, not significant.

Table 2. Association between *EGFR* mutations and CA-repeat in 113 lung adenocarcinoma patients

	EGFR mutation			*P*
	Present		Absent	
	Deletion	Substitution		
Exon 18				
CA-repeat				
<17	0	1 (1.8%)	55 (98.2%)	ns†
≥17	0	2 (3.5%)	55 (96.5%)	
Exon 19				
CA-repeat				
<17	16 (28.6%)	0	40 (71.4%)	0.02
≥17	6 (10.5%)	0	51 (89.5%)	
Exon 21				
CA-repeat				
<17	0	10 (17.9%)	46 (82.1%)	ns
≥17	0	13 (22.8%)	44 (77.2%)	

†Not significant.
EGFR, epidermal growth factor receptor.

with adjustment for gender, age, and smoking status, did not differ significantly by CA-repeat length ($P = 0.19$, Fig. 2a). However, when patients were divided by age category according to the median age (70 years), increased levels of log(*EGFR* mRNA/*β-actin* mRNA) were associated with shorter CA-repeat length in the patients below age 70 ($P = 0.03$, Fig. 2c). Gender and smoking status may also influence the levels of log(*EGFR* mRNA/*β-actin* mRNA) (Fig. 2b,d, respectively). Therefore, we carried out a categorical regression analysis, with adjustment for the confounding variables (age, gender, and smoking status), and found that increased log(*EGFR* mRNA/*β-actin* mRNA) levels were associated with shorter CA-repeat length ($P = 0.02$, Table 3). As for cancer tissues, the median values of EGFR mRNA/β-actin mRNA were 10.7, and 9.8 in longer, and shorter CA-repeat length, respectively. No statistically significant association was found between CA-repeat length and *EGFR* mRNA levels in cancer tissues (data not shown).

Further examining the relationship between CA-repeat length and *EGFR* mutations at exon 19 according to whether log(*EGFR* mRNA/*β-actin* mRNA) was greater than or less than the median (2.10), we found a significantly increased frequency of *EGFR* mutations associated with shorter CA-repeat length only in the patients with higher *EGFR* mRNA expression ($P = 0.045$; Table 4).

Table 3. Categorical regression analysis of log(*EGFR* mRNA/*β-actin* mRNA) in 74 lung adenocarcinoma patients

Variables	β	*P*
Gender (male *versus* female)	−0.08	0.68
Age (<70 *versus* ≥70 years)	−0.12	0.31
Smoking status (non-smokers *versus* smokers)	−0.24	0.24
CA-repeat length (<17 *versus* ≥17)	−0.29	0.02

EGFR, epidermal growth factor receptor.

Logistic regression analysis of *EGFR* mutation frequency at exon 19, taking into account age, gender, and smoking status, also showed a significant contribution of shorter CA-repeat length in the patients with higher *EGFR* mRNA expression ($P = 0.03$, $\beta = -2.6$).

CA-repeat and growth of normal HBE cells. Normal HBE cells with CA 16/15 evidenced faster cell growth than those with CA 20/19 ($P = 0.004$, Fig. 3a). This is consistent with a shorter doubling time for HBE cells with CA 16/15 ($P = 0.006$, Fig. 3b). *EGFR* mRNA levels in HBE cells with CA 16/15 were significantly higher than those in HBE cells with CA 20/19 ($P = 0.017$, Fig. 3c). In addition, EGFR phosphorylation induced by EGF was enhanced in one clone with CA 16/15 and in another clone with CA 8/7, compared with that found in 3 clones with CA 20/19 (Fig. 4). These results suggest an inverse relationship between CA-repeat length and cell growth in HBE cells. Enhanced EGF/EGFR signaling was also found in HBE cells with shorter CA-repeats.

Discussion

EGFR mutations display different clinicopathological features according to the exons at which they occur. The majority of patients with mutations at exon 21 were female non-smokers who were diagnosed with adenocarcinomas showing bronchioloalveolar features, whereas patients with mutations at exon 19 included greater proportions of males and current or former smokers and a smaller proportion with bronchioloalveolar features.[10] *EGFR* mutations at exon 19 were small in-frame deletions, whereas those at exons 18 and 21 were base substitutions. Our findings that *EGFR* mutations were associated with gender, smoking status, and histological type, and occurred often in adenocarcinoma patients but rarely in squamous cell carcinoma patients, are consistent with previous reports.[7–10] Our finding

Table 4. *EGFR* mutations at exon 19 and CA-repeat length in relation to log(*EGFR* mRNA/*β-actin* mRNA) levels

		CA-repeat length		*P**
		<17	≥17	
log(*EGFR* mRNA/*β-actin* mRNA) < median†				
Mutations at exon19	Present	3	2	0.65
	Absent	14	17	
log(*EGFR* mRNA/*β-actin* mRNA) ≥ median				
Mutations at exon19	Present	7	1	0.045
	Absent	13	17	

*The exact *P*-value (two-sided) based on the Pearson χ^2 test.
†Median value of log(*EGFR* mRNA/*β-actin* mRNA) in non-cancerous tissues = 2.10.
EGFR, epidermal growth factor receptor.

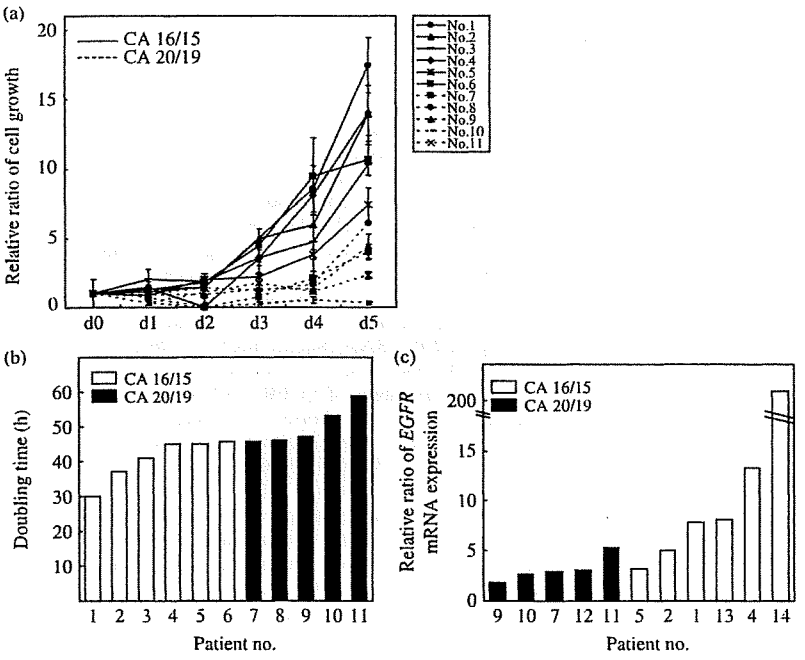

Fig. 3. Growth (a) and doubling time (b) of normal human bronchial epithelial (HBE) cells according to number of CA-repeats. Normal HBE cells were isolated from bronchial mucosal biopsy and subjected to experiments after three passages. Cells (4×10^4 cells/mL) were cultured in 0.5 mL medium, and the number of cells was determined by trypan blue staining after incubation for the indicated period. Relative cell numbers, which were divided by the cell number on day 0, are shown and the values are expressed as means ± SD of triplicate analyzes. (c) Levels of *epidermal growth factor receptor (EGFR)* mRNA in normal HBE cells as determined using real-time reverse transcription-polymerase chain reaction. The relative expression levels of *EGFR* mRNA were determined after correction for *β-actin* mRNA as a control gene. NS, not significant.

doi: 10.1111/j.1349-7006.2008.00804.x

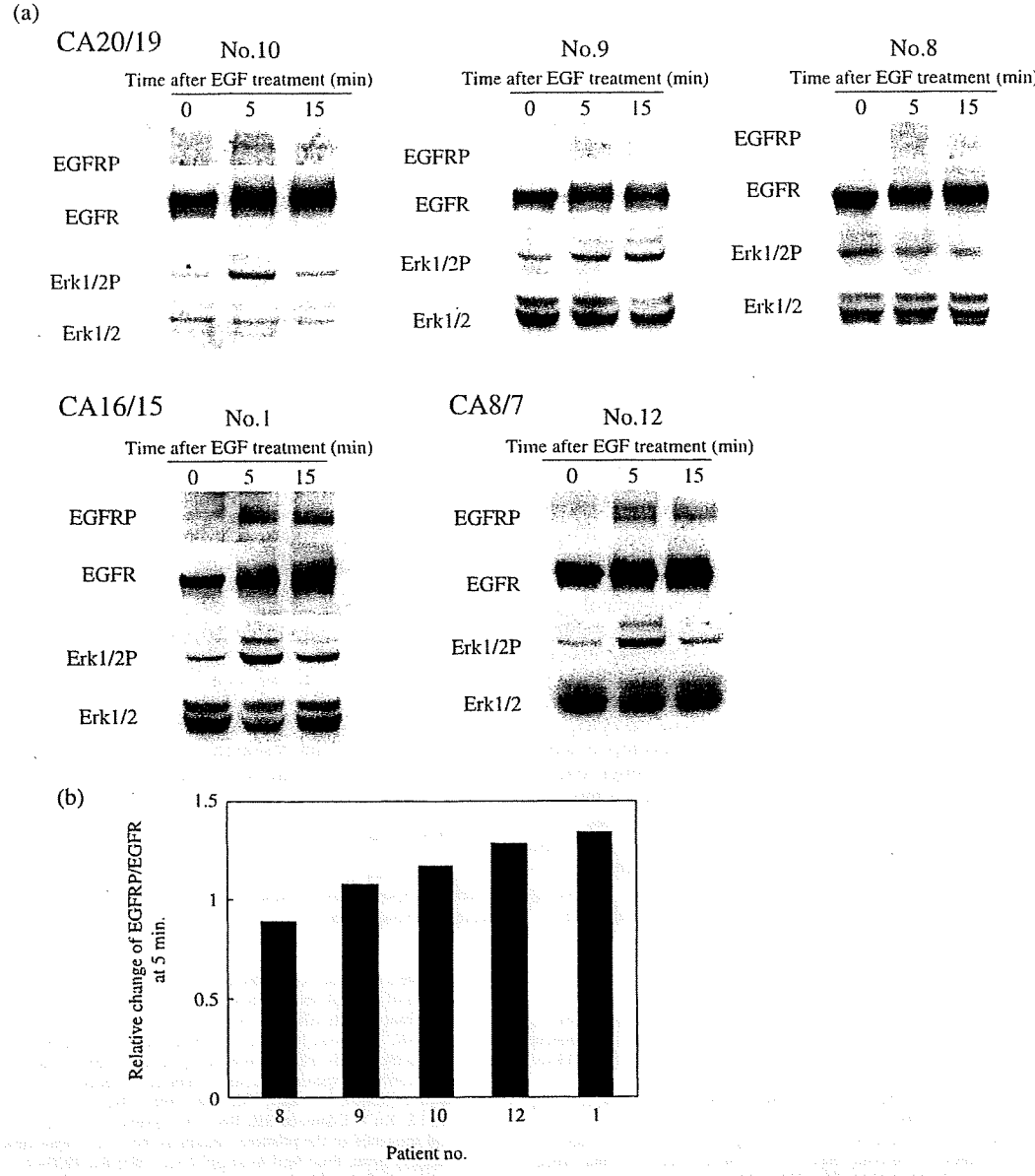

Fig. 4. *Epidermal growth factor receptor (EGFR)* phosphorylation induced by EGF was more apparent in the cells with CA-repeat lengths in the longer/shorter alleles of 16/15 and 8/7 compared to cells with CA-repeat lengths of 20/19. Normal human bronchial epithelial (HBE) cells were first incubated for 24 h in medium without EGF and bovine pituitary extract, then treated with 100 ng/mL EGF for the indicated periods. The numbers of CA-repeats were 20/19 in patients 10, 9, 8, 16/15 in patient 1, and 8/7 in patient 12. Western blot analysis was conducted on whole cell lysates (50 μg) (a). The relative ratio of EGFR phosphorylation was determined as the intensity of EGFR phosphorylation divided with that of EGFR (b).

that CA-repeat length at intron 1 of the *EGFR* gene displayed bimodal distributions in both the shorter and longer alleles is similar to a result reported in non-Chinese Asians.[23]

The results of the present study may be summarized as follows. First, *EGFR* mutations at exon 19 were closely associated with shorter CA-repeat lengths in the shorter allele, but this was not the case with *EGFR* mutations at exons 18 or 21. This implies that CA-repeat length is associated with deletion mutations, but not with substitution mutations. Second, the mechanism relating CA-repeat and *EGFR* mutations at exon 19 was elucidated by our finding that increased intrinsic *EGFR* mRNA expression in non-cancerous tissues was significantly associated with shorter CA-repeats. Third, the relationship between increased *EGFR* mutations and shorter CA-repeat length at exon 19 was apparent in lung adenocarcinoma patients with *EGFR* mRNA expression levels in the upper 50th percentile, but not in patients with lower *EGFR* expression levels. Fourth and lastly, our finding that increased cell growth and enhanced EGF/EGFR signaling occurred in HBE cells with shorter CA-repeat lengths and increased *EGFR* mRNA expression confirms the relationship between CA-repeat length and *EGFR* mRNA expression in the parallel studies mentioned above and provides a clue as to what cell phenotypes might arise from the shorter CA-repeat. Because EGF/EGFR signaling is enhanced in cells undergoing

rapid growth, and we found an inverse relationship between cell growth and CA-repeat length further related to *EGFR* mRNA levels, we conclude that *EGFR* mRNA levels are elevated in cells with shorter CA-repeat lengths, which results in enhanced EGF/EGFR signaling and therefore more rapid cell growth.

As for the mechanism linking CA-repeat length and *EGFR* mRNA expression, it has been postulated that the CA-repeat length at intron 1 influences DNA bendability and hence the binding of repressor protein.[29] However, we are puzzled as to how increased *EGFR* expression could be linked to increased *EGFR* mutations, specifically deletions. Our HBE study demonstrated that shorter CA-repeat lengths are associated with faster cell growth, which might result from increased *EGFR* expression and activation of EGF/EGFR signaling. Further study of the cell phenotypes related to shorter CA-repeat length in association with increased *EGFR* mutations is needed, though, because there are patients with *EGFR* mutations and shorter CA-repeat lengths with no evidence of elevated *EGFR* mRNA expression, which suggests that other mechanisms may be involved.

It has been reported recently that EGFR is also involved in DNA repair. EGFR binds to DNA-dependent protein kinase (DNA-PK) complex and induces the translocation of DNA-PK complex into the nucleus.[30,31] Mutated forms of EGFR abrogate ionizing radiation-induced nuclear EGFR translocation or binding to DNA-PK catalytic subunit (DNA-PKcs), resulting in inhibition of DNA-PK activity.[32] DNA-PK is involved in non-homologous end joining, one of the most important DNA repair systems in mammalians.[33] Because defects in the DNA-PK complex, Ku proteins, and DNA-PKcs enhance tumorigenecity in transgenic mice,[34–36] it is possible that *EGFR* mutations are causally related to cancer development; in fact, transgenic mice carrying the *EGFR* mutations $EGFR^{L858R}$ and $EGFR^{\Delta L747-S752}$ developed adenocarcinomas of the lung.[12] Taken together, these findings lead us to postulate that shorter CA-repeat lengths might enhance intrinsic *EGFR* mRNA expression through altered bendability of the gene, and up-regulated EGFR might cause enhanced cell growth and attenuated DNA repair capacity – specifically against deletions – through physical interaction with DNA-PK complex, resulting in increased *EGFR* mutations in normal lung epithelial cells.

Here we report that an association between *EGFR* mRNA expression and CA-repeat length was found only in non-cancerous tissues. That *EGFR* mRNA expression in cancer tissues was not highly correlated with that in corresponding non-cancerous tissues (correlation coefficient 0.34) is consistent with our finding that cancer tissues did not evidence an association between *EGFR* expression and CA-repeat length. This may be due to the fact that intrinsic *EGFR* expression is altered in cancer cells by various factors including mutations as well as allelic imbalance of *EGFR*, which has also been reported to be associated with CA-repeat length.[25]

As for interethnic differences in CA-repeat length, lengths of less than 17 in Japanese are less frequent than in Caucasians.[23] However, the frequency of *EGFR* mutations is higher in Japanese than in other ethnic groups (primarily Caucasians): 30–50% versus 5–20%.[37–39] Mutations of other genes in lung cancer, specifically *KRAS* and *TP53*, are more frequent in Caucasians than in Japanese. *KRAS* and *TP53* mutations are known to be caused by smoking, and in particular the G:C-to-T:A transversion in *TP53* is generally interpreted as a mutagen footprint.[40] However, *KRAS* and *TP53* mutations occur at relatively high rates even among non-smokers in Japanese (6% and 30%, respectively[7]) and at even higher rates in Caucasians (10% and 47.5%, respectively[40]). *KRAS* and *EGFR* mutations are mutually exclusive, and either type of mutation alone is thought to be sufficient for lung carcinogenesis. These factors relative to mutations in other genes may contribute to interethnic differences in *EGFR* mutation rates as well as different genetic backgrounds.

In future research, we plan to investigate whether the CA-repeat polymorphism can serve as an alternative predictive marker for EGFR tyrosine kinase inhibitor response. We also intend to expand our research on the relationship between CA-repeat and *EGFR* mutations in lung carcinogenesis.

Acknowledgments

This work was supported by grants from the Ministry of Education, Science, Sports and Culture (19591116, Japan), and the Smoking Research Foundation.

References

1 'Cancer Statistics in Japan' Editorial Board. Number of deaths and proportional mortality rates from malignant neoplasms by site in Japan (2005). In: Nomura K, Sobue T, Nakatani H *et al.*, eds. *Cancer Statistics in Japan-2005*. Tokyo: Foundation for Promotion Cancer Research, 2005; 36–9.
2 Mountain CF, Dresler CM. Regional lymph node classification for lung cancer staging. *Chest* 1997; **111**: 1718–23.
3 Lynch TJ, Bell DW, Sordella R *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004; **350**: 2129–39.
4 Paez JG, Janne PA, Lee JC *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004; **304**: 1497–500.
5 Mitsudomi T, Kosaka T, Endoh H *et al.* Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. *J Clin Oncol* 2005; **23**: 2513–20.
6 Shepherd FA, Rosell R, Weighing tumor biology in treatment decisions for patients with non-small cell lung cancer. *J Thorac Oncol* 2007; **2**: S68–76.
7 Kosaka T, Yatabe Y, Endoh H, Kuwano H, Takahashi T, Mitsudomi T. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. *Cancer Res* 2004; **64**: 8919–23.
8 Tokumo M, Toyooka S, Kiura K *et al.* The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers. *Clin Cancer Res* 2005; **11**: 1167–73.
9 Tomizawa Y, Iijima H, Sunaga N *et al.* Clinicopathologic significance of the mutations of the epidermal growth factor receptor gene in patients with non-small cell lung cancer. *Clin Cancer Res* 2005; **11**: 6816–22.
10 Sueoka N, Sato A, Eguchi H *et al.* Mutation profile of EGFR gene detected by denaturing high-performance liquid chromatography in Japanese lung cancer patients. *J Cancer Res Clin Oncol* 2007; **133**: 93–102.
11 Fukuoka M, Yano S, Giaccone G *et al.* Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer (The IDEAL 1 Trial). *J Clin Oncol* 2003; **21**: 2237–46.
12 Politi K, Zakowski MF, Fan PD, Schonfeld EA, Pao W, Varmus HE. Lung adenocarcinomas induced in mice by mutant EGF receptors found in human lung cancers respond to a tyrosine kinase inhibitor or to down-regulation of the receptors. *Genes Dev* 2006; **20**: 1496–510.
13 Ishii S, Xu YH, Stratton RH, Roe BA, Merlino GT, Pastan I. Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. *Proc Natl Acad Sci USA* 1985; **82**: 4920–4.
14 Johnson AC, Ishii S, Jinno Y, Pastan I, Merlino GT. Epidermal growth factor receptor gene promoter. Deletion analysis and identification of nuclear protein binding sites. *J Biol Chem* 1988; **263**: 5693–9.
15 Kageyama R, Merlino GT, Pastan I. Epidermal growth factor (EGF) receptor gene transcription. Requirement for Sp1 and an EGF receptor-specific factor. *J Biol Chem* 1988; **263**: 6329–36.
16 Maekawa T, Imamoto F, Merlino GT, Pastan I, Ishii S. Cooperative function of two separate enhancers of the human epidermal growth factor receptor proto-oncogene. *J Biol Chem* 1989; **264**: 5488–94.
17 Chi DD, Hing AV, Helms C, Steinbrueck T, Mishra SK, Donis-Keller H. Two chromosome 7 dinucleotide repeat polymorphisms at gene loci epidermal growth factor receptor (EGFR) and pro alpha 2 (I): collagen (COL1A2). *Hum Mol Genet* 1992; **1**: 135.
18 Gebhardt F, Zanker KS, Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. *J Biol Chem* 1999; **274**: 13176–80.
19 Zhang W, Weissfeld JL, Romkes M, Land SR, Grandis JR, Siegfried JM. Association of the EGFR intron 1 CA repeat length with lung cancer risk. *Mol Carcinog* 2007; **46**: 372–80.
20 Lee SJ, Kim KM, Chae MH *et al.* No association between dinucleotide repeat polymorphism in intron 1 of the epidermal growth factor receptor gene EGFR and risk of lung cancer. *Cancer Genet Cytogenet* 2007; **172**: 29–32.

doi: 10.1111/j.1349-7006.2008.00804.x

21. Han SW, Jeon YK, Lee KH *et al.* Intron 1 CA dinucleotide repeat polymorphism and mutations of epidermal growth factor receptor and gefitinib responsiveness in non-small-cell lung cancer. *Pharmacogenet Genomics* 2007; **17**: 313–19.
22 Amador ML, Oppenheimer D, Perea S *et al.* An epidermal growth factor receptor intron 1 polymorphism mediates response to epidermal growth factor receptor inhibitors. *Cancer Res* 2004; **64**: 9139–43.
23 Liu W, Innocenti F, Chen P, Das S, Cook EH Jr, Ratain MJ. Interethnic difference in the allelic distribution of human epidermal growth factor receptor intron 1 polymorphism. *Clin Cancer Res* 2003; **9**: 1009–12.
24 Buerger H, Gebhardt F, Schmidt H *et al.* Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression. *Cancer Res* 2000; **60**: 854–7.
25 Nomura M, Shigematsu H, Li L *et al.* Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. *Plos Med* 2007; **4**: 715–27.
26 Taron M, Ichinose Y, Rosell R *et al.* Activating mutations in the tyrosine kinase domain of the epidermal growth factor receptor are associated with improved survival in gefitinib-treated chemorefractory lung adenocarcinomas. *Clin Cancer Res* 2005; **11**: 5878–85.
27 Xu XC, Lee JJ, Wu TT, Hoque A, Ajani JA, Lippman SM. Increased retinoic acid receptor-beta4 correlates in vivo with reduced retinoic acid receptor-beta2 in esophageal squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 826–9.
28 Sueoka N, Lee HY, Walsh GL, Hong WK, Kurie JM. Posttranslational mechanisms contribute to the suppression of specific cyclin: CDK complexes by all-trans retinoic acid in human bronchial epithelial cells. *Cancer Res* 1999; **59**: 3838–44.
29 Brandt B, Meyer-Staeckling S, Schmidt H, Agelopoulos K, Buerger H. Mechanisms of egfr gene transcription modulation: relationship to cancer risk and therapy response. *Clin Cancer Res* 2006; **12**: 7252–60.
30 Bandyopadhyay D, Mandal M, Adam L, Mendelsohn J, Kumar R. Physical interaction between epidermal growth factor receptor and DNA-dependent protein kinase in mammalian cells. *J Biol Chem* 1998; **273**: 1568–73.
31 Dittmann K, Mayer C, Fehrenbacher B *et al.* Radiation-induced epidermal growth factor receptor nuclear import is linked to activation of DNA-dependent protein kinase. *J Biol Chem* 2005; **280**: 31182–9.
32 Das AK, Chen BP, Story MD, Sato M, Minna JD, Chen DJ, Nirodi CS. Somatic mutations in the tyrosine kinase domain of epidermal growth factor receptor (EGFR) abrogate EGFR-mediated radioprotection in non-small cell lung carcinoma. *Cancer Res* 2007; **67**: 5267–74.
33 Smith GC, Jackson SP. The DNA-dependent protein kinase. *Genes Dev* 1999; **13**: 916–34.
34 Mondello C, Rebuzzini P, Dolzan M, Edmonson S, Taccioli GE, Giulotto E. Increased gene amplification in immortal rodent cells deficient for the DNA-dependent protein kinase catalytic subunit. *Cancer Res* 2001; **61**: 4520–5.
35 Espejel S, Martin M, Klatt P, Martin-Caballero J, Flores JM, Blasco MA. Shorter telomeres, accelerated ageing and increased lymphoma in DNA-PKcs-deficient mice. *EMBO Rep* 2004; **5**: 503–9.
36 Difilippantonio MJ, Zhu J, Chen HT *et al.* DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* 2000; **404**: 510–14.
37 Yang SH, Mechanic LE, Yang P *et al.* Mutations in the tyrosine kinase domain of the epidermal growth factor receptor in non-small cell lung cancer. *Clin Cancer Res* 2005; **11**: 2106–10.
38 Marchetti A, Martella C, Felicioni L *et al.* EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. *J Clin Oncol* 2005; **23**: 857–65.
39 Cappuzzo F, Hirsch FR, Rossi E *et al.* Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst* 2005; **97**: 643–55.
40 Le Calvez F, Mukeria A, Hunt JD *et al.* TP53 and KRAS mutation load and types in lung cancers in relation to tobacco smoke: distinct patterns in never, former, and current smokers. *Cancer Res* 2005; **65**: 5076–83.

Stroke
JOURNAL OF THE AMERICAN HEART ASSOCIATION

American Stroke Association
A Division of American Heart Association

Cigarette Smoking, Systolic Blood Pressure, and Cardiovascular Diseases in the Asia-Pacific Region

Koshi Nakamura, Federica Barzi, Tai-Hing Lam, Rachel Huxley, Valery L. Feigin, Hirotsugu Ueshima, Jean Woo, Dongfeng Gu, Takayoshi Ohkubo, Carlene M.M. Lawes, Il Suh, Mark Woodward and for the Asia Pacific Cohort Studies Collaboration

Stroke 2008;39;1694-1702; originally published online Mar 6, 2008;
DOI: 10.1161/STROKEAHA.107.496752
Stroke is published by the American Heart Association. 7272 Greenville Avenue, Dallas, TX 72514
 Print ISSN: 0039-2499. Online ISSN: 1524-4628

Cigarette Smoking, Systolic Blood Pressure, and Cardiovascular Diseases in the Asia-Pacific Region

Koshi Nakamura, MD; Federica Barzi, PhD; Tai-Hing Lam, MD; Rachel Huxley, DPhil; Valery L. Feigin, MD; Hirotsugu Ueshima, MD; Jean Woo, MD; Dongfeng Gu, MD; Takayoshi Ohkubo, MD; Carlene M.M. Lawes, PhD; Il Suh, MD; Mark Woodward, PhD; for the Asia Pacific Cohort Studies Collaboration

Background and Purpose—Smoking and increased levels of blood pressure (BP) substantially increase the risk of cardiovascular diseases (CVD). If these 2 risk factors have a synergistic impact on cardiovascular events, lowering BP and quitting smoking will contribute more to reducing CVD than would be expected from ignoring their interaction.

Methods—Individual participant data were combined from 41 cohorts, involving 563 144 participants (82% Asian). During a median of 6.8 years follow-up, 4344 coronary heart disease (CHD) and 5906 stroke events were recorded. Repeat measures of systolic blood pressure (SBP) were used to adjust for regression dilution bias. Hazard ratios (HRs) and 95% confidence intervals (CIs) for SBP by cigarette smoking status were estimated from Cox proportional hazard models adjusted for age and stratified by study and sex.

Results—Data suggested a log-linear relationship between SBP and all subtypes of CVD. The HRs relating SBP to both CHD and ischemic stroke were broadly similar irrespective of smoking status ($P \geq 0.1$). For hemorrhagic stroke (intracerebral hemorrhage), the HRs (95% CIs) for an additional 10 mm Hg increment in SBP were 1.81 (1.73 to 1.90) for present smokers and 1.66 (1.59 to 1.73) for nonsmokers ($P=0.003$). For every subtype of cardiovascular events, similar results were found for analyses involving only fatal events.

Conclusions—Smoking exacerbated the impact of SBP on the risk of hemorrhagic stroke. Although quitting smoking and lowering BP are both crucial for prevention of CVD, combining the 2 could be expected to have extra beneficial effect on preventing hemorrhagic stroke. **(*Stroke*. 2008;39:1694-1702.)**

Key Words: smoking ■ blood pressure ■ cardiovascular diseases ■ coronary heart disease ■ stroke

Nonoptimal levels of blood pressure (BP) and smoking are the first and second most common causes of death in the world, and, together, these 2 risk factors account for more than 20% of the global burden of premature death.[1,2] In particular, increased BP[3–7] and smoking[7–11] are major risk factors for cardiovascular diseases (CVD), including coronary heart disease (CHD) and stroke. Previous studies have indicated that smoking and increased BP interact to increase markers of cardiovascular risk, including levels of plasma fibrinogen[12] and carotid intima-media thickness.[13] Hence, a combination of raised BP and smoking may have a synergistic impact on cardiovascular events, especially those caused by atherosclerosis and thrombosis.[14] If such an interaction exists, multifactorial interventions aimed at both lowering BP and quitting smoking will contribute more to reducing CVD than expected from past data where their interaction has not been quantified.

Several epidemiological studies have examined the combined effects of nonoptimal levels of BP and smoking on cardiovascular events.[11,15–23] Some studies, at least partially, observed a synergistic effect between BP and smoking status for the risk of CVD,[15] CHD,[16–18] and stroke (predominately ischemic),[11,16,19,20] whereas other studies did not observe any such effect.[21,22] The majority of these studies were based on small datasets and crude classifications of BP and smoking status, and few examined the possible interaction effect between BP and smoking status for each subtype of CVD. For hemorrhagic stroke, only 1 case-control study[23] examined the interaction between BP and smoking status; it reported that interaction was present. Overall, however, the question as to

Received June 14, 2007; final revision received September 13, 2007; accepted October 26, 2007.

From the Nutrition and Lifestyle Division (K.N., F.B., R.H.), The George Institute for International Health, Sydney, Australia; Department of Community Medicine (T.H.L.), University of Hong Kong, People's Republic of China; Clinical Trials Research Unit (V.L.F., C.M.M.L.), University of Auckland, New Zealand; Department of Health Science (H.U.), Shiga University of Medical Science, Otsu, Japan; Division of Geriatrics (J.W.), Department of Medicine & Therapeutics, Chinese University of Hong Kong, People's Republic of China; Cardiovascular Institute and Fu Wai Hospital (D.G.), Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People's Republic of China; Department of Planning for Drug Development and Clinical Evaluation (T.O.), Tohoku University Graduate School of Pharmaceutical Science and Medicine, Sendai, Japan; Department of Preventive Medicine (I.S.), Yonsei University College of Medicine, Seoul, Korea; Mount Sinai Medical Center (M.W.), New York.

Correspondence to Koshi Nakamura, MD, Nutrition and Lifestyle Division, The George Institute for International Health, PO Box M201, Missenden Road, Camperdown, NSW 2050, Australia. E-mail knakamura@george.org.au

Stroke is available at http://stroke.ahajournals.org

DOI: 10.1161/STROKEAHA.107.496752

whether such an interaction exists, and the nature of this interaction (synergistic or otherwise), remains unresolved. The aim of the present study was to examine this issue using data from the Asia Pacific Cohort Studies Collaboration (APCSC); an individual participant data overview of prospective cohort studies conducted in the Asia-Pacific region. The large size of the dataset provides an ideal opportunity to explore the joint associations of risk factors with cardiovascular events. In particular, the large numbers of both hemorrhagic and ischemic stroke events makes it possible to measure the risk for each subtype of stroke reliably. Additionally, APCSC provides a unique opportunity to compare the association of risk factors with cardiovascular events between Asian populations and the "Western" populations of Australia and New Zealand.

Methods

Participating Studies

Details of APCSC are described elsewhere.[24,25] Briefly, APCSC is an overview of preexisting cohort studies in the Asia-Pacific region which had at least 5000 person-years of follow-up and recorded age, sex, and BP at baseline, and vital status at the end of the follow-up. Studies were excluded from APCSC if enrolment was dependent on having a particular condition or risk factor. Additionally, for analyses in this report, only persons aged ≥20 years at study entry with information on both BP and smoking status were included.

Measurement of Baseline Variables

In most studies, BP was measured at rest in the seated position using a standard mercury sphygmomanometer. Cigarette smoking habit was self-reported at study baseline. All studies included here recorded present smoking status (present smoker or not). Some studies additionally recorded whether individuals were present, former, or never smokers, and some recorded cigarettes per day for smokers. Because most studies, including APCSC, have demonstrated that the association between systolic blood pressure (SBP) and cardiovascular events is stronger than that of other BP indices in most age and gender groups,[26,27] we analyzed data on SBP in this report. Cohorts were classified as Asian if the participants were recruited from mainland China, Hong Kong, Japan, Korea, Singapore, Taiwan, or Thailand and as ANZ if the participants were from Australia or New Zealand. This classification largely represented a split by ethnicity into Asians and Whites.

Outcomes

All studies reported deaths by underlying cause; a subset of studies also reported nonfatal cardiovascular events. Outcomes were classified according to the Ninth Revision of the International Classification of Diseases (ICD-9). Outcomes in this report, including fatal and nonfatal events, were CHD (ICD-9: 410 to 414) and stroke (430 to 438), divided into hemorrhagic stroke (intracerebral hemorrhage; 431.0 to 432.9), ischemic stroke (433.0 to 434.9), and other strokes. Because most studies identified events using record linkage, verification of pathological types of stroke was not routinely reported. All data provided to the Secretariat were checked for completeness and consistency and recoded, when necessary, to maximize comparability across cohorts. Summary reports were referred back to principal investigators of each collaborating study for review and confirmation.

Statistical Methods

Cox proportional hazard regression models adjusted by age and stratified by study and sex[28] were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for SBP by smoking status (nonsmokers, including former smokers, and present smokers). To determine the associations between "usual" level of SBP and the outcomes of interest, estimates were adjusted for regression dilution bias.[3,29] Repeat measurements of SBP on up to 7 occasions, between 2 and 20 years after the baseline measurement, were obtained from 16 studies for a total of 67 210 participants. These repeat measures were used to estimate a regression dilution attenuation coefficient for SBP (1.9), using a linear mixed regression model that accounted for the heterogeneity of variance between studies and within-subject correlation.[30] Log-linearity of the associations between SBP and each subtype of cardiovascular event was explored by categorical analyses in which participants were classified into 4 groups according to levels of baseline SBP (<130, 130 to 144, 145 to 159, and ≥160 mm Hg) chosen so as to have approximately equal numbers of all cardiovascular events across the groups. Corresponding 95% CIs were calculated by the "floating absolute risk method."[29] HRs and 95% CIs were also derived for a 10 mm Hg increase in the level of SBP. The interaction effect between SBP and smoking status was assessed using likelihood ratio tests comparing the models with main effects only with the models that included the interaction term.[29] In addition to analyses of the overall APCSC, predefined subgroup analyses were performed by sex, region (Asia and ANZ), and age at risk (<65 and ≥65 years).[24]

Further analyses were conducted on subsamples of the total population which had more detailed information on smoking status. In one of the subsamples, participants were classified as "present" if they smoked currently, "never smokers" if they had never smoked, and "former smokers" if they had smoked but reported having already quit at study baseline. HRs for a 10-mm Hg increase in the level of SBP were estimated for each group by this smoking status and compared using similar methods to the main analyses. Similarly, dose-response analyses were done on the subset where both the mean number of cigarettes smoked per day and never smoking were recorded, comparing never smokers with <20 and ≥20 cigarettes per day for present smokers. Groups of <20 and ≥20 were chosen to provide an approximately equal partition; 20 cigarettes corresponds to 1 standard pack.

Results

Characteristics of the Study Population

Information on SBP and smoking status was available from 41 cohorts (93% of all studies in APCSC); 32 from Asia (Table 1). Overall, 563 144 participants were included in the analysis (82% Asians; 35% female) with a mean age of 47 years. Over one-third (37%) of study participants were classified as present smokers at baseline, but the prevalence of smoking differed by sex and region: in Asia, 59% of men and 5% of women were present smokers versus 20% and 14%, respectively, in ANZ. In Asia, mean age and SBP were similar between smokers and nonsmokers (45 years versus 45 years and 122 mm Hg versus 121 mm Hg), but in ANZ, present smokers were both younger and had a lower SBP than nonsmokers: 48 years versus 54 years and 133 mm Hg versus 138 mm Hg, respectively. These mean values of age and SBP were weighted, rather than crude, averages across studies.

Information on former smoking status was available from 34 cohorts (24 in Asia). In these, 63 941 (13%) of participants were former smokers, 261 319 (51%) were never smokers, and 187 416 (37%) were present smokers. In Asian cohorts, 15% of men and 22% of women who had ever smoked had quit, compared to 68% and 59%, respectively, in ANZ. Of these 34 cohorts, 24 also recorded information on the average number of cigarettes smoked per day. Among the 97 540 present smokers in these cohorts, 44% consumed 20 cigarettes or more per day: in Asia, 44% for men and 21% for women, versus 52% and 43%, respectively, in ANZ.

Table 1. Study Population Characteristics by Smoking Status at Baseline

	Nonsmokers				Current Smokers			
Study Name	n	Age (years) mean (SD)	SBP (mm Hg) mean (SD)	Female (%)	n	Age (years) mean (SD)	SBP (mm Hg) mean (SD)	Female (%)
Akabane	1321	55 (8)	125 (19)	77	513	53 (7)	124 (19)	2
Anzhen	5992	54 (13)	129 (24)	69	2386	53 (12)	130 (22)	20
Anzhen02	3287	47 (8)	122 (18)	64	864	46 (8)	122 (17)	1
Beijing aging	1472	70 (9)	143 (25)	62	620	69 (8)	137 (25)	24
Capital Iron Steel Company	1367	45 (8)	125 (19)	0	3775	45 (8)	123 (19)	0
CISCH	1576	44 (7)	117 (17)	69	591	45 (8)	122 (16)	2
Civil service workers	5739	47 (5)	125 (18)	47	3501	47 (5)	126 (18)	10
CVDFACTS	4455	47 (15)	118 (19)	70	1274	48 (16)	119 (18)	4
East Beijing	806	45 (15)	125 (23)	64	322	41 (15)	124 (21)	20
EGAT	1980	43 (5)	121 (17)	38	1514	43 (5)	121 (16)	3
Fangshan	1591	47 (10)	136 (26)	86	1028	48 (10)	135 (25)	36
Guangzhou occupational	87 400	41 (6)	115 (15)	41	79 295	42 (7)	116 (14)	1
Hisayama	918	57 (12)	135 (26)	82	683	55 (10)	135 (26)	22
Hong Kong	2428	79 (7)	150 (25)	63	555	77 (6)	148 (24)	33
Kinmen	1824	63 (10)	138 (23)	64	721	64 (9)	136 (21)	9
KMIC	98 631	44 (7)	121 (14)	54	61 611	45 (7)	125 (14)	0
Konan	857	52 (16)	130 (20)	75	369	51 (16)	130 (18)	9
Miyama	756	61 (10)	134 (22)	73	317	60 (9)	130 (22)	13
Ohasama	1793	60 (11)	127 (17)	78	447	58 (12)	132 (18)	7
Saitama	2588	54 (12)	135 (20)	80	1027	55 (12)	136 (19)	17
Seven cities cohorts	7019	54 (12)	130 (25)	70	3792	54 (12)	129 (23)	26
Shanghai factory workers	5198	47 (7)	124 (21)	51	4149	50 (7)	126 (23)	5
Shibata	1573	57 (11)	130 (21)	82	777	57 (11)	133 (20)	8
Shigaraki town	2657	58 (14)	132 (19)	77	1073	56 (14)	132 (20)	16
Shirakawa	3023	48 (12)	127 (22)	79	1617	48 (12)	126 (21)	8
Singapore heart	1807	40 (13)	124 (22)	61	514	41 (14)	122 (18)	7
Singapore NHS92	2699	39 (12)	119 (19)	62	606	39 (12)	118 (17)	8
Six cohorts	10 465	44 (7)	119 (18)	76	8922	45 (7)	119 (17)	12
Tanno/Soubetsu	1214	51 (7)	134 (20)	78	764	51 (7)	132 (21)	14
Tianjin	4586	56 (13)	139 (28)	64	4749	54 (11)	134 (25)	39
Xi'an	1020	44 (6)	126 (21)	49	675	45 (6)	125 (20)	10
Yunnan	2138	58 (10)	126 (22)	9	4443	55 (9)	123 (21)	0
Total Asia	270 180	45 (10)	121 (18)	53	193 494	45 (9)	122 (17)	4
ALSA	1486	78 (6)	148 (22)	48	124	76 (6)	148 (26)	48
ANHF	7043	44 (14)	126 (18)	53	2234	41 (13)	125 (18)	45
Busselton	5155	45 (17)	138 (25)	59	2634	44 (16)	137 (24)	37
Canberra	728	77 (5)	145 (21)	46	93	76 (5)	147 (22)	39
Fletcher challenge	7899	46 (15)	127 (17)	30	2427	40 (13)	124 (15)	22
Melbourne	36 630	55 (9)	138 (20)	60	4655	53 (8)	135 (19)	47
Newcastle	4567	52 (11)	133 (20)	53	1362	50 (10)	131 (20)	40
Perth	7625	46 (13)	130 (20)	51	2605	43 (13)	129 (19)	40
WAAAAS	10 870	72 (4)	157 (21)	0	1333	71 (4)	157 (22)	0
Total ANZ	82 003	54 (14)	138 (22)	47	17 467	48 (15)	133 (22)	37
Total	352 183	47 (12)	125 (20)	52	210 961	45 (9)	123 (18)	7

SD indicates standard deviation; SBP, systolic blood pressure; ANZ, Australia and New Zealand; ALSA, Australian Longitudinal Study of Aging; ANHF, Australian National Heart Foundation; CISCH, Capital Iron and Steel Company Hospital; EGAT, Electricity Generating Authority of Thailand; KMIC, Korean Medical Insurance Corporation; NHS92, National Health Study 1992; WAAAAS, Western Australian AAA Screenees.

Cardiovascular Outcomes

In total, there were 3 907 543 person-years of follow-up; the median follow-up was 6.8 years (6.8 years for present smokers and 6.7 years for nonsmokers) but, for both present smokers and nonsmokers, it was shorter in Asia (6.8 years versus 6.0 years) than in ANZ (8.3 years versus 8.2 years; Table 2). In addition to information on fatal events available from all cohorts, data on nonfatal CHD events were available from 14 studies and on nonfatal strokes from 12 studies. During follow-up, 4344 CHD (1569 in Asia) and 5906 stroke (4218 in Asia) fatal and nonfatal events were recorded: 76% (n=3282) of CHD events were fatal. Over 80% of CHD events were myocardial infarction. Of all stroke events, 2001 (1550 in Asia) were classified as ischemic and 1645 (1441 in Asia) as hemorrhagic: 30% (n=608) of ischemic stroke and 73% (n=1207) of stroke events were fatal. Diagnosis of ischemic or hemorrhagic stroke was documented by CT/MRI/autopsy investigations in 56% of fatal and 65% of nonfatal, strokes. The percentage of CHD among all CVD (CHD plus stroke) was similar between smokers and nonsmokers (40% versus 44%): these percentages in ANZ (61% versus 64%) were more than double those in Asia (29% versus 25%). The percentage of hemorrhagic strokes among all strokes was similar between smokers and nonsmokers (30% versus 26%): these percentages were higher in Asia (34% versus 34%) than in ANZ (13% versus 12%).

The Association Between SBP and CHD by Smoking Status

The HR for CHD increased log-linearly with higher levels of SBP in both smokers and nonsmokers (Figure 1A). The HRs (95% CIs) comparing the top to the bottom group of SBP were 2.27 (2.05 to 2.52) for present smokers and 2.20 (2.05 to 2.36) for nonsmokers. The HR for a 10-mm Hg increase in SBP level was also similar for present smokers and nonsmokers (Figure 2): 1.29 (1.24 to 1.34) and 1.24 (1.21 to 1.28), respectively (probability value for interaction=0.14). The coronary HRs for present smokers and nonsmokers were similar in all sex, age, and region subgroups. Similar results (not shown) were found for analyses involving fatal events only.

In the subsample of studies for which information on former smokers was available, the HRs for CHD associated with a 10-mm Hg increase in SBP were similar for present smokers and never smokers. However, the HR was lower in former smokers than in present or never smokers: 1.28 (1.22 to 1.33) for present smokers, 1.14 (1.09 to 1.20) for former smokers, and 1.30 (1.25 to 1.35) for never smokers (probability value for interaction=0.0001). In the subsample of studies with information on cigarette consumption, the HRs for CHD tended to increase with increasing consumption of cigarettes: 1.27 (1.21 to 1.32) for never smokers, 1.30 (1.19 to 1.43) for <20 cigarettes per day, and 1.41 (1.28 to 1.54) for ≥20 cigarettes per day (probability value for interaction=0.11).

The Association Between SBP and Ischemic Stroke by Smoking Status

Similar to CHD, there was no evidence of an interaction between BP and smoking for risk of ischemic stroke: the HR for ischemic stroke increased log-linearly with higher levels of SBP in both present smokers and nonsmokers (Figure 1B). The HRs (95% CIs) comparing the highest with the lowest group of SBP were 3.71 (3.22 to 4.27) for present smokers and 3.82 (3.43 to 4.26) for nonsmokers. The HR for a 10-mm Hg increase in SBP level was similar for present smokers and nonsmokers in all subgroups (Figure 2). Overall HRs (95% CIs) were 1.50 (1.43 to 1.57) for present smokers and 1.47 (1.41 to 1.53) for nonsmokers (probability value for interaction=0.53). Similar results (not shown) were found for analyses involving fatal events only.

In the subsample with information on former smokers, the HR for a 10-mm Hg increase in SBP was similar for present smokers, former smokers, and never smokers: 1.44 (1.36 to 1.52), 1.41 (1.29 to 1.53), and 1.41 (1.34 to 1.49), respectively (probability value for interaction=0.86). Among those participants with information on cigarettes per day there was marginally nonsignificant evidence of an increasing effect of SBP with increasing cigarette consumption. The HRs were 1.30 (1.20 to 1.41) for never smokers, 1.47 (1.26 to 1.70) for <20 cigarettes per day, and 1.62 (1.34 to 1.97) for ≥20 cigarettes per day (probability value for interaction=0.06).

The Association Between SBP and Hemorrhagic Stroke by Smoking Status

The HR for hemorrhagic stroke increased with higher levels of SBP in both present smokers and nonsmokers (Figure 1C). There was evidence to support a synergistic effect of smoking on the association between SBP and hemorrhagic stroke risk: the HRs (95% CIs) for hemorrhagic stroke comparing the group with the highest to that with the lowest SBP values were 9.32 (8.15 to 10.67) for present smokers and 7.05 (6.27 to 7.92) for nonsmokers. The excess risk of hemorrhagic stroke associated with a 10-mm Hg higher SBP level increased in present smokers compared with nonsmokers by 15 percentage points (ie, 81% versus 66%) (Figure 2): 1.81 (1.73 to 1.90) versus 1.66 (1.59 to 1.73); probability value for interaction=0.003. Subgroup analysis found indications of this synergistic effect in most subgroups, although it was statistically significant only for men (P=0.01), in Asian study centers (P=0.05), and individuals aged 65 years or over (P=0.008) (Figure 2). Restricting the analysis to fatal hemorrhagic events resulted in a similar pattern: HR (95% CI) for a 10-mm Hg increase in SBP was 1.82 (1.72 to 1.92) for present smokers and 1.67 (1.59 to 1.75) for nonsmokers (probability value for interaction=0.01).

The HR for a 10-mm Hg increase in SBP was higher in present smokers than in former smokers and never smokers: 1.87 (1.77 to 1.97) versus 1.55 (1.40 to 1.71) and 1.68 (1.58 to 1.78), respectively (probability value for interaction=0.0008). In the subsample with information on cigarettes per day, the HRs increased with higher dose of smoking: 1.60 (1.47 to 1.75) for never smokers, 1.85 (1.65 to 2.08) for <20 cigarettes per day, and 1.95 (1.72 to 2.22) for ≥20 cigarettes per day (probability value for interaction=0.01).

A sensitivity analysis using only data from participants (n=126 956) in which information on the use of antihypertensive medication status at study baseline was available indicated that further adjustment for use of antihypertensive

Table 2. Fatal and Nonfatal Cardiovascular Events by Smoking Status

	Nonsmokers					Current Smokers				
			Stroke					Stroke		
Study Name	Median FUP	CHD	Isch	Hem	Others	Median FUP	CHD	Isch	Hem	Others
Akabane	11.0	15	9	5	11	11.0	13	7		6
Anzhen	4.3	50	74		7	4.3	15	32	20	3
Anzhen02	3.0		11	43		3.0	1	3	1	1
Beijing aging	4.8				61	4.8				25
Capital Iron Steel Company	12.5	13	15	20		12.5	70	77	45	9
CISCH	3.3	9			6	3.3	5			3
Civil service workers	6.7					6.7	1		1	1
CVDFACTS	6.1	10	6	5	10	5.8	3	1	3	4
East Beijing	16.0	12	10	8	2	17.4	8	4	3	1
EGAT	11.4	9			8	11.4	24			8
Fangshan	3.6	2	15	6	4	3.6	3	5	2	2
Guangzhou occupational	7.3	60		68	37	7.2	106		99	58
Hisayama	25.1	40	129	29	19	22.6	49	101	39	11
Hong Kong	2.5	73	5	14	41	2.5	13	1	2	10
Kinmen	2.9	6			8	2.9	4			6
KMIC	4.0	107	187	161	150	4.0	171	245	164	147
Konan	6.4		6	2	2	6.4	2	1	1	
Miyama	6.6	1	2		2	6.6	1	4	1	1
Ohasama	4.1	2	21	9	4	4.1	5	16	2	2
Saitama	11.0	14	19	9	10	10.0	10	8	6	3
Seven cities cohorts	2.7	51	66	109	6	2.7	33	51	73	2
Shanghai factory workers	14.0	33			114	14.0	53			141
Shibata	20.0	40	46	23	62	20.0	27	31	13	34
Shigaraki town	4.4	2	2	2	1	4.4	1	2		6
Shirakawa	17.5	29	18	20	12	17.5	36	21	11	5
Singapore heart	14.7	40	16	6	37	14.2	26	6	1	9
Singapore NHS92	6.2	22	11	1	19	6.2	11	3	3	8
Six cohorts	9.0	6	33	50	7	8.3	41	71	41	6
Tanno/Soubetsu	16.4	8	7	7	5	16.4	16	3	9	2
Tianjin	6.1	65	58	97	43	6.1	49	64	90	22
Xi'an	19.7	12	8	17	2	19.7	23	7	7	
Yunnan	4.5	7	5	42		4.5	11	7	51	1
Total Asia	6.0	738	779	753	690	6.8	831	771	688	537
ALSA	4.7	77	7	8	34	3.3	4			3
ANHF	8.4	55	1		10	8.3	22	1		5
Busselton	26.5	767	153	57	407	26.5	480	85	40	207
Canberra	9.6	106	5	4	23	8.4	14	1	1	4
Fletcher challenge	5.7	202	56	7	101	5.8	71	11	2	17
Melbourne	8.5	262	10	28	43	8.7	61	1	7	11
Newcastle	8.5	78	3	6	15	9.4	59		3	7
Perth	14.4	127	3	7	29	14.4	68	1	3	20
WAAAAS	3.2	285	98	29	86	3.2	37	15	2	11
Total ANZ	8.2	1959	336	146	748	8.3	816	115	58	285
Total	6.7	2697	1115	899	1438	6.8	1647	886	746	822

FUP indicates follow-up (years); CHD, coronary heart disease; Isch, ischemic; Hem, hemorrhagic; ANZ, Australia and New Zealand; ALSA, Australian Longitudinal Study of Aging; ANHF, Australian National Heart Foundation; CISCH, Capital Iron and Steel Company Hospital; EGAT, Electricity Generating Authority of Thailand; KMIC, Korean Medical Insurance Corporation; NHS92, National Health Study 1992; WAAAAS, Western Australian AAA Screenees; Blanks indicate that the event was not reported for that study.

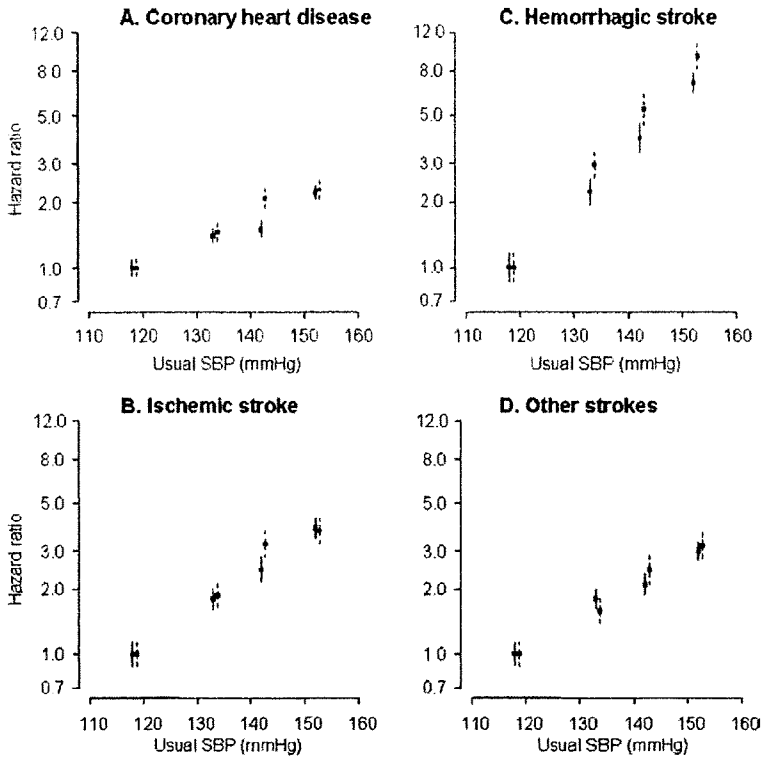

Figure 1. Associations between usual systolic blood pressure (SBP) and overall events by smoking status for: (A) coronary heart disease, (B) ischemic stroke, (C) hemorrhagic stroke, and (D) other strokes. The hazard ratio (95% confidence interval) for the lowest group of SBP is fixed at 1.0, separately for present smokers and nonsmokers. Analyses are adjusted by age and stratified by study and sex. The dashed (right) and continuous (left) lines represent present smokers and nonsmokers, respectively. (Probability values for log-linearity <0.0001 for all.)

medication did not attenuate the difference in risk estimates between present smokers and nonsmokers. The HR (95% CI) for a 10-mm Hg increase in SBP was 1.42 (1.24 to 1.63) for present smokers and 1.23 (1.11 to 1.36) for nonsmokers, after age adjustment (probability value for interaction=0.09), and 1.39 (1.21 to 1.60) and 1.20 (1.09 to 1.33), respectively, after age and use of antihypertensive medication adjustment (probability value for interaction=0.08).

The Association Between SBP and Other Strokes by Smoking Status

For completeness, Figure 1D shows the categorical analyses for other strokes. As with ischemic and hemorrhagic strokes, the HR increased with higher levels of SBP in both present smokers and nonsmokers. The HRs (95% CIs) comparing the highest with the lowest group of SBP were 3.17 (2.76 to 3.64) for present smokers and 3.01 (2.76 to 3.64) for nonsmokers. The HRs (95% CIs) for a 10-mm Hg increment in SBP level were 1.40 (1.33 to 1.47) in present smokers and 1.36 (1.31 to 1.41) in nonsmokers (probability value for interaction=0.33).

Discussion

The present study demonstrates a log-linear relationship of SBP with every subtype of CVD, for both smokers and nonsmokers, with no evidence of a threshold effect down to usual levels of SBP of 115 mm Hg. For hemorrhagic stroke, there was evidence that SBP and smoking have a synergistic effect such that smoking increases the excess risk associated with a 10-mm Hg increment in SBP by about 15 percentage points. Our data suggest that this interaction may be specific to men and older participants, but is unlikely to be specific to region because of the marginal differences between smokers and nonsmokers in both Asia and ANZ apparent from Figure 2. By comparison, the excess relative risk associated with increments in SBP for both CHD and ischemic stroke was broadly similar for smokers and nonsmokers.

The prevailing cause of CHD and ischemic stroke is occlusion of the coronary and cerebral arteries due to atherosclerosis and thrombosis.[14] Some previous reports suggest that nonoptimal levels of BP combined with smoking may promote atherothrombogenesis.[12,13] Kiyohara and colleagues[16] observed an interaction effect between BP and smoking status for CHD in women but not in men, and 1 study[17] observed such an effect in women. Meanwhile, 1 study[18] observed such an effect in men. In a case-control study, Ohgren and colleagues[19] reported an interaction effect between BP and smoking status for all strokes (78% of which were ischemic). Two Japanese studies[11,16] observed such a potentiation for ischemic stroke among men (but not women[16]), as did the British Regional Heart Study,[20] in which the majority of strokes would be expected to be ischemic in origin. By contrast, 2 studies[21,22] in populations where ischemic stroke predominates did not observe such a potentiation for all strokes. These null findings are consistent with our results based on the simple assessment of present smoking status (ie, present/nonsmokers, and present/former/never smokers), suggesting that smoking does not exacerbate the association between SBP and the risk of CHD and ischemic stroke. Furthermore, as most of the previous studies used a relatively crude classification of smoking and hypertensive status, previous positive findings of an interaction may have been attributable to chance alone. There was however some sug-

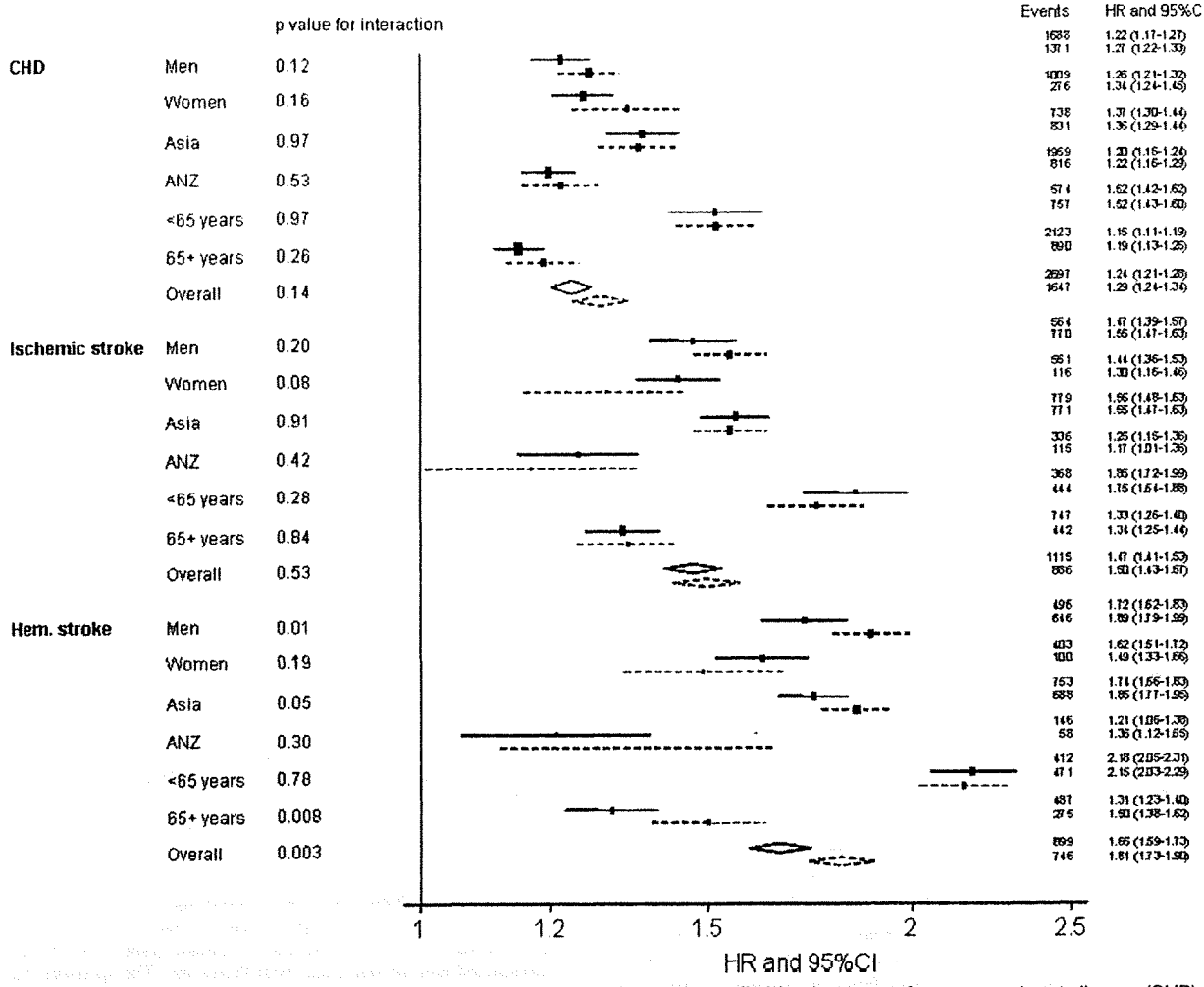

Figure 2. Hazard ratios (HRs) associated with a 10-mm Hg increase in usual systolic blood pressure for coronary heart disease (CHD), ischemic stroke, and hemorrhagic (Hem) stroke, in present smokers and nonsmokers, by sex, region, age, and overall. Analyses are adjusted by age and stratified by study and sex. The horizontal lines (or widths of diamonds for overall results) show 95% confidence intervals (CIs). The probability values shown are for the test of interaction between systolic blood pressure and smoking status. The dashed (lower) and continuous (upper) lines represent present smokers and nonsmokers, respectively.

gestion of an interaction for CHD and ischemic stroke when restricting the present analysis to those studies with information on cigarette consumption, in agreement with an earlier study[11] which reported that the risk of ischemic stroke increased more strongly with higher dose of smoking among individuals with hypertension compared with those without. By contrast, another study[21] reported that the smoking dose-related risk for all strokes was similar for both those with and without hypertension.

Unlike CHD and ischemic stroke, the prevailing cause of hemorrhagic stroke is rupture resulting from fragility (including microaneurysms) of the intracerebral penetrating arteries caused by nonoptimal levels of BP or amyloid angiopathy.[31,32] This accounts for the stronger association between BP and hemorrhagic stroke risk compared with CHD, although the risk related to increased levels of BP is similar for ischemic and hemorrhagic stroke.[4,5] By contrast, the excess risk attributable to smoking for hemorrhagic stroke is less than it is for either CHD or ischemic stroke.[9–11] As regards a pathophysiological mechanism behind the interaction for hemorrhagic stroke observed in the present study, we can only speculate that smoking may promote the weakening of the intracranial blood vessels caused by high levels of BP or amyloid angiopathy. Only Thrift and colleagues[23] have examined the interaction between BP and smoking status for hemorrhagic stroke events. In this case-control study, a significant synergistic interaction was observed only in men, which is consistent with our findings. The sex-specific effect that we observed may have been a chance finding as a consequence of the few events among the smaller population of female smokers (n=14 031), compared with male smokers (n=196 930). The regional specificity may result from the difficulty in observing hemorrhagic stroke events due to a much smaller number participants and a lower event rate of hemorrhagic stroke in ANZ (204 events per 99 470 ANZ participants) compared with Asia (1411 events per 463 674

Asian participants). However, neither of these explanations would explain the age-specific significant effect, wherein the interaction only occurred among those aged 65 years or over: 883 events for <65 years and 762 events for ≥65 years.

The present study has some limitations. First, some cohorts in APCSC do not have information on other risk factors for CVD at baseline, restricting our ability to adjust for important covariates which may explain the observed interaction effects between BP and smoking. Serum total cholesterol, which is positively associated with CHD and ischemic stroke events, and inversely with hemorrhagic stroke events,[33] was available on 353 158 individuals; data on other potentially useful covariates was less common. However, adjustment for total cholesterol had negligible impact on any of the reported results (not shown). Second, we had limited data on daily dose of smoking and little information on how smoking status changed during follow-up, which did not allow any reliable analyses of follow-up smoking status comparable to our treatment of SBP. Third, the main analysis was not adjusted for antihypertensive medication status because of a lack of this information for more than 70% of participants, although the sensitivity analysis suggests that it may have little material impact on the results. Finally, there was lack of standardization of methods and procedures among the participating studies in APCSC, because the participating studies were originally independent of each other. For instance, only 56% of fatal and 65% of nonfatal strokes were objectively (using CT/MRI or autopsy findings) classified as ischemic or hemorrhagic in origin. The Hisayama study in Japan,[34] 1 of the APCSC participating studies, investigated the accuracy of diagnosis of each subtype of CVD using autopsies in the 1960s, 1970s, and 1980s. The accuracy of diagnosis was similar for ischemic and hemorrhagic stroke (confirmation rate 60% to 70%), which was better than the accuracy for CHD (46%).[34] Therefore, misclassification of stroke subtype may have introduced bias the extent of which would have varied across the studies.

In conclusion, we have shown that a combination of present smoking and nonoptimal levels of BP appears to have a synergistic impact on the risk of hemorrhagic stroke, at least among men and in the elderly, although the underlying pathophysiological mechanism is unclear, and we cannot exclude that similar synergism may occur among younger people and women. Furthermore, we cannot affirm the absence of interaction between BP and smoking for CHD and ischemic stroke. Further studies allowing for better verification of pathological types of stroke, better assessment of smoking status and other variables, and using a larger and more standardized dataset, are warranted to determine whether the interaction between BP and smoking really exists for each subtype of CVD, what mechanism explains the interaction, and how specific it is to demographic groups. Although quitting smoking and lowering BP are both crucial for prevention of CVD, combining the two could be expected to have extra beneficial effect on preventing hemorrhagic stroke. Thus, smoking cessation initiatives should be targeted more rigorously for hypertensive patients to prevent hemorrhagic stroke.

Appendix

The Asia Pacific Cohort Studies Collaboration

Executive Committee

M. Woodward (Chair). X. Fang, D.F. Gu, R Huxley, Y. Imai, T.H. Lam, W.H. Pan, A. Rodgers, I. Suh, H.C. Kim, H. Ueshima,

Participating Studies and Principal Collaborators

Aito Town: A Okayama, H Ueshima, H Maegawa; *Akabane*: N Aoki, M Nakamura, N Kubo, T Yamada; *Anzhen 02*: ZS Wu; *Anzhen*: CH Yao, ZS Wu; *Australian Longitudinal Study of Aging*: Mary Luszcz; *Australian National Heart Foundation*: TA Welborn; *Beijing Aging*: Z Tang; *Beijing Steelworkers*: LS Liu, JX Xie; *Blood Donors' Health*: R Norton, S Ameratunga, S MacMahon, G Whitlock; *Busselton*: MW Knuiman; *Canberra-Queanbeyan*: H Christensen; *Capital Iron and Steel Company Hospital Cohort (CISCH)*: J Zhou, XH Yu; *Capital Iron and Steel Company*: XG Wu; *Civil Service Workers*: A Tamakoshi; *CVDFACTS*: WH Pan; *Electricity Generating Authority of Thailand (EGAT)*: P Sritara; *East Beijing*: ZL Wu, LQ Chen, GL Shan; *Fangshan Farmers*: DF Gu, XF Duan; *Fletcher Challenge*: S MacMahon, R Norton, G Whitlock, R Jackson; *Guangzhou*: YH Li; *Guangzhou Occupational*: TH Lam, CQ Jiang; *Hisayama*: Y Kiyohara, H Arima, M Iida; *Hong Kong*: J Woo, SC Ho; *Huashan*: Z Hong, MS Huang, B Zhou; *Kinmen*: JL Fuh; *Kounan Town*: H Ueshima, Y Kita, SR Choudhury; *Korean Medical Insurance Corporation*: I Suh, SH Jee, IS Kim; *Melbourne Cohort*: G Giles; *Miyama*: T Hashimoto, K Sakata; *Newcastle*: A Dobson; *Ohasama*: Y Imai, T Ohkubo, A Hozawa; *Perth*: K Jamrozik, M Hobbs, R Broadhurst; *Saitama*: K Nakachi; *Seven Cities*: XH Fang, SC Li, QD Yang; *Shanghai Factory Workers*: ZM Chen; *Shibata*: H Tanaka; *Shigaraki*: Y Kita, A Nozaki, H Ueshima; *Shirakawa*: H Horibe, Y Matsutani, M Kagaya; *Singapore Heart*: K Hughes, J Lee; *Singapore 92*: D Heng, SK Chew; *Six Cohorts*: BF Zhou, HY Zhang; *Tanno/Soubetsu*: K Shimamoto, S Saitoh; *Tianjin*: ZZ Li, HY Zhang; *Western Australian AAA Screenees*: P Norman, K Jamrozik; *Xi'an*: Y He, TH Lam; *Yunnan*: SX Yao.

Sources of Funding

This project has received support from a National Health and Medical Research Council of Australia program grant (358395) and an unrestricted educational grant from Pfizer Inc. The sponsors had no influence on design, analysis, or interpretation of results, and took no part in the writing of this paper. C.M.M. Lawes is supported by a National Heart Foundation (New Zealand) Fellowship.

Disclosures

None.

References

1. Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJ. Comparative Risk Assessment Collaborating Group. Selected major risk factors and global and regional burden of disease. *Lancet.* 2002;360: 1347–1360.
2. World Health Organisation. *The World Health Report 2002. Reducing risks, promoting healthy life.* Geneva: World Health Organisation; 2002.
3. MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyer A, Stamler J. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet.* 1990;335:765–774.
4. Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet.* 2002;360:1903–1913 (errata: *Lancet.* 2003;361:1060).
5. Asia Pacific Cohort Studies Collaboration. Blood pressure and cardiovascular disease in the Asia Pacific region. *J Hypertens.* 2003;21: 707–716.
6. Nippon Data 80 Research Group. Impact of elevated blood pressure on mortality from all causes, cardiovascular diseases, heart disease and stroke among Japanese: 14 year follow-up of randomly selected popu-

lation from Japanese – Nippon data 80. *J Hum Hypertens.* 2003;17: 851–857.
7. Ariesen MJ, Claus SP, Rinkel GJ, Algra A. Risk factors for intracerebral hemorrhage in the general population: a systematic review. *Stroke.* 2003; 34:2060–2065.
8. Kuller LH, Ockene JK, Meilahn E, Wentworth DN, Svendsen KH, Neaton JD. Cigarette smoking and mortality. MRFIT Research Group. *Prev Med.* 1991;20:638–654.
9. Asia Pacific Cohort Studies Collaboration. Smoking, quitting, and the risk of cardiovascular disease among women and men in the Asia-Pacific region. *Int J Epidemiol.* 2005;34:1036–1045.
10. Ueshima H, Choudhury SR, Okayama A, Hayakawa T, Kita Y, Kadowaki T, Okamura T, Minowa M, Iimura O. Cigarette smoking as a risk factor for stroke death in Japan: NIPPON DATA80. *Stroke.* 2004;35: 1836–1841.
11. Yamagishi K, Iso H, Kitamura A, Sankai T, Tanigawa T, Naito Y, Sato S, Imano H, Ohira T, Shimamoto T. Smoking raises the risk of total and ischemic strokes in hypertensive men. *Hypertens Res.* 2003;26:209–217.
12. Tuut M, Hense HW. Smoking, other risk factors and fibrinogen levels: evidence of effect modification. *Ann Epidemiol.* 2001;11:232–238.
13. Howard G, Wagenknecht LE, Burke GL, Diez-Roux A, Evans GW, McGovern P, Nieto FJ, Tell GS. Cigarette smoking and progression of atherosclerosis: The Atherosclerosis Risk in Communities (ARIC) Study. *JAMA.* 1998;279:119–124.
14. Labarthe DR. *Epidemiology and Prevention of Cardiovascular Diseases: A Global Challenge.* Gaithersburg, MD: Aspen Publication Inc; 1998.
15. Khalili P, Nilsson PM, Nilsson JA, Berglund G. Smoking as a modifier of the systolic blood pressure-induced risk of cardiovascular events and mortality: a population-based prospective study of middle-aged men. *J Hypertens.* 2002;20:1759–1764.
16. Kiyohara Y, Ueda K, Fujishima M. Smoking and cardiovascular disease in the general population in Japan. *J Hypertens.* 1990;8(Suppl 5):S9–S15.
17. Janzon E, Hedblad B, Berglund G, Engstrom G. Tobacco and myocardial infarction in middle-aged women: a study of factors modifying the risk. *J Intern Med.* 2004;256:111–118.
18. Stamler J, Wentworth D, Neaton JD. Is relationship between serum cholesterol and risk of premature death from coronary heart disease continuous and graded? Findings in 356,222 primary screenees of the Multiple Risk Factor Intervention Trial (MRFIT). *JAMA.* 1986;256: 2823–2828.
19. Ohgren B, Weinehall L, Stegmayr B, Boman K, Hallmans G, Wall S. What else adds to hypertension in predicting stroke? An incident case-referent study. *J Intern Med.* 2000;248:475–482.
20. Shaper AG, Phillips AN, Pocock SJ, Walker M, Macfarlane PW. Risk factors for stroke in middle aged British men. *BMJ.* 1991;302:1111–1115.
21. Colditz GA, Bonita R, Stampfer MJ, Willett WC, Rosner B, Speizer FE, Hennekens CH. Cigarette smoking and risk of stroke in middle-aged women. *N Engl J Med.* 1988;318:937–941.
22. Wolf PA, D'Agostino RB, Kannel WB, Bonita R, Belanger AJ. Cigarette smoking as a risk factor for stroke. The Framingham Study. *JAMA.* 1988;259:1025–1029.
23. Thrift AG, McNeil JJ, Donnan GA. The risk of intracerebral haemorrhage with smoking. The Melbourne Risk Factor Study Group. *Cerebrovasc Dis.* 1999;9:34–39.
24. Asia Pacific Cohort Studies Collaboration. Determinants of cardiovascular disease in the Asia Pacific region: protocol for a collaborative overview of cohort studies. *CVD Prevention.* 1999;2:281–289.
25. Woodward M, Barzi F, Martiniuk A, Fang X, Gu DF, Imai Y, Lam TH, Pan WH, Rodgers A, Suh I, Sun HJ, Ueshima H, Huxley R. Cohort profile: The Asia Pacific Cohort Studies Collaboration. *Int J Epidemiol.* 2006;35:1412–1416.
26. Asia Pacific Cohort Studies Collaboration. Blood pressure indices and cardiovascular disease in the Asia Pacific region: a pooled analysis. *Hypertension.* 2003;42:69–75.
27. Sesso HD, Stampfer MJ, Rosner B, Hennekens CH, Gaziano JM, Manson JE, Glynn RJ. Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in Men. *Hypertension.* 2000;36:801–807.
28. Cox DR. Regression models and life tables (with discussion). *J Royal Stat Soc.* 1972;34:187.
29. Woodward M. *Epidemiology: Study Design and Data Analysis.* II edn. Boca Raton: Chapman and Hall/CRC; 2005.
30. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol.* 1990; 132:734–745.
31. Ferro JM. Update on intracerebral haemorrhage. *J Neurol.* 2006;253: 985–999.
32. Russell RW. How does blood-pressure cause stroke? *Lancet.* 1975;306: 1283–1285.
33. Asia Pacific Cohort Studies Collaboration. Cholesterol, coronary heart disease, and stroke in the Asia Pacific region. *Int J Epidemiol.* 2003;32: 563–572.
34. Hasuo Y, Ueda K, Kiyohara Y, Wada J, Kawano H, Kato I, Yanai T, Fujii I, Omae T, Fujishima M. Accuracy of diagnosis on death certificates for underlying causes of death in a long-term autopsy-based population study in Hisayama, Japan; with special reference to cardiovascular diseases. *J Clin Epidemiol.* 1989;42:577–584.

J Cancer Res Clin Oncol (2008) 134:211–217
DOI 10.1007/s00432-007-0272-3

ORIGINAL PAPER

The association between RAD18 Arg302Gln polymorphism and the risk of human non-small-cell lung cancer

Hirotaka Kanzaki · Mamoru Ouchida · Hiroko Hanafusa · Hiromasa Yamamoto · Hiromitsu Suzuki · Masaaki Yano · Motoi Aoe · Kazue Imai · Hiroshi Date · Kei Nakachi · Kenji Shimizu

Received: 10 April 2007 / Accepted: 22 June 2007 / Published online: 12 July 2007

Abstract
Purpose The repair enzyme RAD18 plays a key role in the post-replication repair process in various organisms from yeast to human, and the molecular function of the RAD18 protein has been elucidated. Single nucleotide polymorphism (SNP) of arginine (Arg, CGA) or glutamine (Gln, CAA) at codon 302 is known on RAD18; however, the association between the SNP and the risk of any human cancers including non-small-cell lung cancer (NSCLC) has not been reported. We therefore investigated the relationship between the polymorphism and the development and progression of human NSCLC.
Methods The study population included 159 patients with NSCLC and 200 healthy controls. The SNP was genotyped by polymerase chain reaction with the confronting two-pair primer (PCR-CTPP) assay. Genotype frequencies were compared between patients and controls, and the association of genotypes with clinicopathological parameters was also studied.
Results The Gln/Gln genotype was significantly more frequent in NSCLC patients (20.7%) than in healthy controls (11.5%)($P = 0.003$). The increased risk was detected in NSCLC patients with the Gln/Gln genotype [Odds ratio (OR) = 2.63, 95% confidence interval (CI)=1.38–4.98]. As to the relationship of the SNP with clinicopathological parameters of NSCLC, significantly higher risks were detected in lung squamous cell carcinoma (LSC) (OR = 4.40, 95% CI = 1.60–12.1).
Conclusions Our results suggested that Gln/Gln genotype of the RAD18 SNP has the increased risk of NSCLC, especially of LSC. This is the first report to provide evidence for an association between the RAD18 Arg302Gln polymorphism and human NSCLC risk.

Keywords SNPs · RAD18 · Non-small-cell lung cancer (NSCLC) · Cancer predisposition

Abbreviations

LAD	Lung adenocarcinoma
LSC	Lung squamous cell carcinoma
NSCLC	Non-small-cell lung cancer
OR	Odds ratio
PCNA	Proliferating cell nuclear antigen
PCR-CTPP	Polymerase chain reaction with the confronting two-pair primers
PRR	Postreplication repair
SNP	Single nucleotide polymorphism

H. Kanzaki · M. Ouchida (✉) · H. Hanafusa · K. Shimizu
Department of Molecular Genetics,
Graduate School of Medicine,
Dentistry and Pharmaceutical Sciences,
Okayama University, 2-5-1 Shikata-cho,
Okayama 700-8558, Japan
e-mail: ouchidam@md.okayama-u.ac.jp

H. Yamamoto · H. Suzuki · M. Yano · M. Aoe · H. Date
Department of Cancer and Thoracic Surgery,
Graduate School of Medicine,
Dentistry and Pharmaceutical Sciences,
Okayama University, Okayama 700-8558, Japan

K. Imai · K. Nakachi
Department of Radiobiology/Molecular Epidemiology,
Radiation Effects Research Foundation,
Hiroshima 732-0815, Japan

Introduction

DNA in living cells is damaged by environmental damaging agents and mutagens, such as UV light and mutagenic chemicals (Hoeijmakers 2001). DNA damage must be repaired by

DNA repair systems. However, when the DNA repair systems are stalled or saturated, and such DNA damages are thus not removed before the onset of DNA replication, single-stranded gaps are generated. These gaps will be filled by the postreplication repair (PRR) system. The RAD6 pathway is known to be central to PRR (Lawrence 1994) and RAD6 epistasis group proteins, such as RAD5, RAD18, RAD30, MMS2 and UBC13, are all involved in the pathway. In this pathway, RAD18 and RAD6 are two of the most important proteins and play a key role. RAD18 is a single-strand DNA binding protein with a RING finger domain, and has ubiquitin-ligating enzymes (E3) activity (Joazeiro and Weissman 2000). RAD6 is an ubiquitin-conjugating enzyme (E2) in the proteasome protein degradation system (Sung et al. 1990, 1991b; Wood et al. 2003). RAD18 forms a tight complex with RAD6 (Bailly et al. 1994, 1997a; b). Although RAD6 interacts with several ubiquitin-ligating enzymes (E3), the interaction with RAD18 is essential for carrying out PRR (Wood et al. 2003; Bailly et al. 1994; Dohmen et al. 1991; Sung et al. 1991a).

RAD18 knockout cells of mouse embryonic stem cells (Tateishi et al. 2003) and of chicken DT40 cells (Yamashita et al. 2002) were hypersensitive to various DNA-damaging agents and showed defective PRR. Genomic instability of these cells was demonstrated by increased rates of the sister chromatid exchange and integration of exogenous DNA (Tateishi et al. 2003; Yamashita et al. 2002). RAD18 contributes to the maintenance of genomic stability through PRR and dysfunction of RAD18 increases the frequency of homologous recombination as well as illegitimate recombination (Shekhar et al. 2002). Furthermore, dysfunction of RAD18 is thought to lead to the development of cancer (Friedberg 2003).

The genetic polymorphisms of DNA repair genes have been analyzed to determine susceptibility to several cancers, including lung (Ito et al. 2004; Ryk et al. 2006), colorectal (Yamamoto et al. 2005), breast (Costa et al. 2006), head and neck (Huang et al. 2005), bladder cancer (Zhu et al. 2007) and leukemia (Bolufer et al. 2006). The *RAD18* gene is known to have a single nucleotide polymorphism (SNP) at codon 302, encoding either arginine (Arg, CGA) or glutamine (Gln, CAA), as known as rs#373572 in the dbSNP; NCBI Reference SNP (refSNP) Cluster Report. In the present study, we found a significant correlation of the SNP with NSCLC. This is, to our knowledge, the first report providing evidence for an association between the RAD18 Arg302Gln polymorphism and human NSCLC risk.

Materials and methods

Subjects

We studied frozen specimens of 159 cases stored at −80°C obtained from Japanese patients with primary NSCLC treated by curative intent surgical resection in Okayama University Hospital (Okayama, Japan), after acquiring informed consent from each patient, between 1994 and 2003. The case groups consisted of 105 lung adenocarcinomas (LAD), 48 lung squamous-cell carcinomas (LSC), 3 adeno/squamous-cell carcinomas and 3 large cell carcinomas (107 men, 52 women; mean age 66.2 years). The clinical stage and pathological grade in most patients were confirmed by operation and pathology. The clinical staging and histological classification of cancers were defined according to the criteria of UICC Tumor-Node-Metastasis Classification of Malignant Tumors (TNM), sixth edition, 2002, (ICD-O C34 for lung). For the controls, each of the 200 healthy controls we analyzed was selected by computer-aided randomization among five individuals matched in smoking habit, gender and age (within 5 years) for each lung cancer patient, all of which were from the subjects of cohort studies on a Japanese general population older than 40 years of age in a town near the Saitama Cancer Center. A population of this town has increased because of a population influx from other areas, with a social increase rate of about 5% every year for 15 years. Informed consent was obtained from all cases and controls concerned. This study was approved by The Bioethics Committee of Okayama University Medical School.

DNA extraction

Genomic DNA of 159 patients was isolated from the noncancerous region of the resected specimens or from the mononuclear cells of the peripheral blood using SDS/proteinase K treatment, phenol-chloroform extraction and ethanol precipitation. Genomic DNA of 200 healthy controls was extracted from peripheral lymphocytes.

Genetic analysis

Genotyping of the RAD18 Arg302Gln polymorphism was carried out by polymerase chain reaction using the confronting two-pair primer (PCR-CTPP) technique (Hamajima et al. 2000; Hamajima 2001). According to the sequence of the human *RAD18* gene shown in database, we designed two sets of paired primers. The first set of primers was as follows: forward primer 1, 5′-ATA CCC ATC ACC CAT CTT C-3′ and reverse primer 1, 5′-GTC TTC TCT ATA TTT TCG ATT TCT T-3′ for the A (Gln) allele amplifying a 146 bp band. The second set of primers was as follows: forward primer 2, 5′-TTA ACA GCT GCT GAA ATA GTT CG-3′ and reverse primer 2, 5′-CTG AAA TAG CCC ATT AAC ATA CA-3′ for the G (Arg) allele amplifying a 106 bp band. A 206 bp band was designed between the forward primer 1 and the reverse primer 2. Genomic DNA (20 ng) was assessed in 20 μl of

reaction mixture containing 40 μM of each dNTP, 1X PCR buffer, 8 pmol of the forward primer 1 and reverse primer 2, 24 pmol of the forward primer 2 and reverse primer 1 and 0.5 unit of the Taq DNA polymerase (Takara, Kyoto, Japan). The PCR amplification was initiated by a denaturing step at 94°C for 3 min, followed by 35 cycles at 94°C for 30 s, 64°C for 1 min, 72°C for 1 min, and a final extension step at 72°C for 7 min. For genotyping, the PCR products were subjected to electrophoresis in 3% agarose gel with ethidium bromide staining and then visualized on a UV transilluminator. The allele types were determined as follows; 205 and 106 bp for the G/G (Arg/Arg) genotype, 205 and 146 bp for the A/A (Gln/Gln) genotype and 205, 146 and 106 bp for the G/A (Arg/Gln) genotype. In order to confirm the allele types, some PCR products were processed with the Big Dye terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA), then analyzed and confirmed on an ABI 3100 sequencer (Applied Biosystems).

Statistical analysis

We compared the allele frequencies of the polymorphism in the *RAD18* gene between NSCLC patient group and healthy control group. The distribution of the *RAD18* genotype (Arg/Arg, Arg/Gln, Gln/Gln) in all of the patients and the controls was tested for adherence to the Hardy–Weinberg equilibrium. The Chi-square test was used to compare the genotype distribution between patients and controls. The odds ratio (OR) and 95% confidence interval (95% CI) were used to estimate the risk of association with genotype. The OR and 95% CI was adjusted for age, gender and smoking habit by an unconditional logistic regression model using the SPSS software Ver.12.0 (SPSS Inc., Tokyo, Japan).

Results

Assessment of cancer risk by RAD18 genotyping

The characteristics of the 159 NSCLC patients and the 200 healthy controls are shown in Table 1. There were no significant differences in gender, age or smoking status between these two groups. Pack-year equivalents were used for smoking status (however, we could not obtain the smoking status for 5 of 159 NSCLC patients).

The representative PCR-CTPP patterns and sequence patterns were shown in Fig. 1a, b, respectively. Significant differences in the genotype frequency were evident between NSCLC patients and controls (Table 2). The frequencies of Arg/Arg, Arg/Gln and Gln/Gln genotype were found to be 29.6, 49.7 and 20.7% in the NSCLC patients and 43.0, 45.5 and 11.5% in the controls, respectively. All of the results fitted the Hardy–Weinberg equilibrium. In comparison to Arg/Arg genotype, the most significantly increased risk was found in NSCLC patients with Gln/Gln genotype with an adjusted OR of 2.57 (95% CI, 1.35–4.89). Thus, this result suggested that the homozygous Gln/Gln genotype has an increased risk of NSCLC.

Table 1 Characteristics of NSCLC patients and healthy controls

	Patients	Controls	*P*-value
	n (%) (*n* = 159)	*n* (%) (*n* = 200)	
Gender			0.874[b]
Male	107 (67.3)	133 (66.5)	
Female	52 (32.7)	67 (33.5)	
Age (years ± SD)[a]	66.2 ± 9.94	65.6 ± 9.42	
Smoking habit			0.909[c]
No-smoker	50 (31.4)	63 (31.5)	
Smoker	104 (65.4)	137 (68.5)	
<20 pack-years	5 (4.8)	17 (12.4)	
≥20 pack-years	97 (92.3)	87 (63.5)	
Unknown	2 (2.9)	33 (24.1)	
Unknown	5 (3.2)	0 (0.0)	

[a] Age shows the mean age of each group with standard deviation

[b] *P*-values were for the differences in the number of males and females between patients and controls and were calculated by Chi-square test

[c] *P*-values were for the differences in the number of smokers and non-smokers between patients and controls and were calculated by Chi-square test

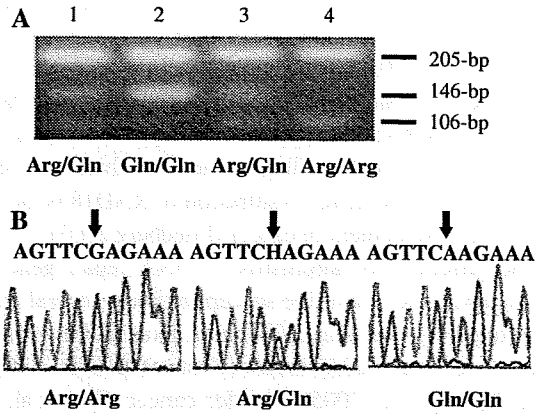

Fig. 1 The single nucleotide polymorphism at codon 302 of the *RAD18* gene. **a** The PCR-CTPP patterns of the *RAD18* SNP. The PCR product was electrophoresed in 3% agarose gel. Two fragments of 205 and 106 bp show the G/G (Arg/Arg) genotype, two fragments of 205 and 146 bp show the A/A (Gln/Gln) genotype, and three fragments of 205-, 146 and 106 bp show the G/A (Arg/Gln) genotype. The case number and genotypes are shown at the *top* and *bottom*, respectively. **b** The direct sequence patterns of the *RAD18* SNP. The SNP, Arg (CGA) or Gln (CAA), is indicated by an *arrow* above the sequence

Table 2 The *RAD18* genotypes in patients and controls

RAD18	Patients	Controls	*P*-value	OR (95% CI)	
Genotype	*N* (%)	*N* (%)		Crude	Adjusted[b]
Arg/Arg	47 (29.6)	86 (43.0)		1 (Reference)	1 (Reference)
Arg/Gln	79 (49.7)	91 (45.5)	0.051[a]	1.59 (1.00–2.53)	1.60 (1.00–2.56)
Gln/Gln	33 (20.7)	23 (11.5)	0.003[a]	2.63 (1.38–4.98)	2.57 (1.35–4.89)
Total	159	200			
Allele frequencies			0.002		
Arg	173 (54.4)	263 (67.8)			
Gln	145 (45.6)	137 (34.2)			

[a] *P*-values were calculated for the difference in genotype frequencies against Arg/Arg by Chi-square test

[b] ORs were adjusted for age, gender and smoking status. Patients whose smoking status was not known were excluded when ORs were calculated

The association between the *RAD18* genotype and clinicopathological features

We next analyzed the relationship between the genotype distribution and the clinicopathological parameters. Strong association between the risk of lung squamous-cell carcinoma (LSC) and genotype distribution was shown in Table 3. The adjusted OR of LSC patients with Gln/Gln genotype was 4.40 (95% CI, 1.60–12.1), whereas the same genotype exhibited a marginal risk for lung adenocarcinoma (LAD) with a borderline significance (adjusted OR = 1.97, 95% CI, 0.94–4.12). Differentiated grade, TNM classification, gender and smoking habit were not associated with the frequency of genotype or allele (Table 4).

Discussion

In the present study, we examined whether the SNP at codon 302 in the *RAD18* gene is associated with the risk for development of NSCLC, and found significant differences in the genotype distribution between the NSCLC patients and the healthy controls. Our findings suggest that this SNP is associated with the development of the NSCLC, and the susceptibility to the NSCLC is enhanced by the Gln/Gln genotype. However, this SNP does not appear to be associated with progression or metastasis of the NSCLC, as the *RAD18* genotype showed no correlation with the clinicopathological characteristics, except histological types. The Gln/Gln genotype was detected more frequently in the NSCLC patients, and the individuals with the Gln/Gln genotype showed a 2.6-fold higher risk of NSCLC. Furthermore, as for the LSC patients, a strong association between the Gln/Gln genotype and the development risk was detected (OR = 4.40, 95% CI = 1.60–12.1). Notably, the heterozygotes (Arg/Gln) exhibited an intermediate risk, still with statistic significance, for both whole NSCLC (OR = 1.60, 95% CI = 1.00–2.56) and LSC (OR = 2.40, 95% CI = 1.09–5.29), indicating a dose-response effect of the Gln allele. This shows that the Gln allele may be defined as the responsive risk-allele. It would be of great interest to see the effects of the SNP on incidence of NSCLC in Europeans and Africans, since the frequency of the individuals with the Gln/Gln genotype is much higher (60%) in these races than in Asian people (8–18%)(rs#373572 in the dbSNP). Giving the high risk of the Gln/Gln genotype for LSC among NSCLC, the ethnic difference may well explain, at least in part, the higher proportion of LSC among NSCLC in Caucasians than in Asians.

Table 3 Association between the RAD18 genotype distribution and histological cell type of patients

Characteristics	Genotype (%)				OR[a] (95% CI)	
	Arg/Arg	*Arg/Gln*	*Gln/Gln*	Total	*Arg/Gln*	*Gln/Gln*
Controls	86 (43.0)	91 (45.5)	23 (11.5)	200		
All patients	47 (29.6)	79 (49.7)	33 (20.7)	159	1.60 (1.00–2.56)	2.57 (1.35–4.89)
LAD	34 (32.4)	53 (50.5)	18 (17.1)	105	1.51 (0.89–2.56)	1.97 (0.94–4.12)
LSC	11 (22.9)	25 (52.1)	12 (25.0)	48	2.40 (1.09–5.29)	4.40 (1.60–12.1)
Others	2 (33.3)	1 (16.7)	3(50.0)	6		

LAD lung adenocarcinoma, *LSC* lung squamous-cell carcinoma

[a] ORs were adjusted for age, gender and smoking status. The Arg/Arg genotype of healthy controls was defined as the reference