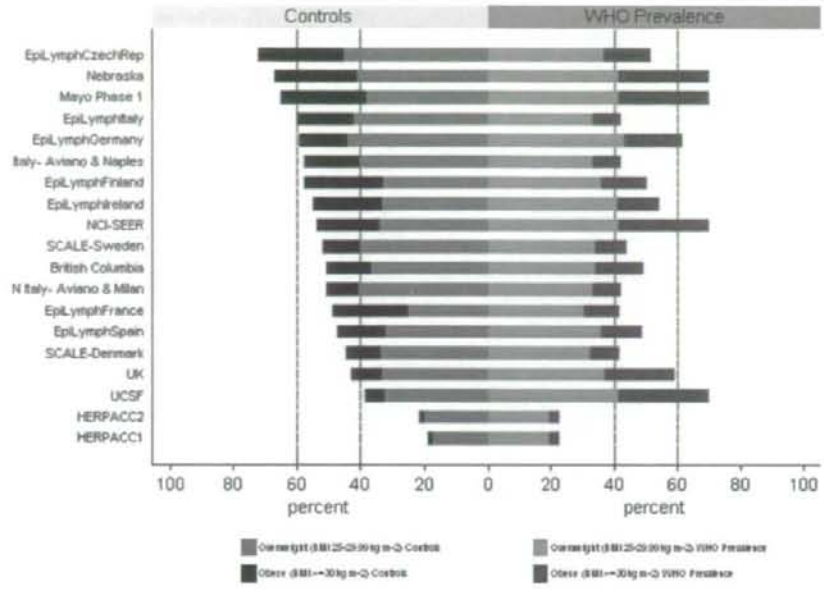

FIGURE 4 – Comparison of Control and WHO overweight/obesity prevalences by study. Overweight (BMI 25–29.99 kg m^{-2}) and Obesity (BMI $\geq$ 30 kg m^{-2}) prevalence from the World Health Organisation (WHO) Global Database on Body Mass Index (http://www.who.int/bmi/). WHO prevalence was derived from the most recent published age- and sex- standardized BMI data calculated from height and weight measured in clearly defined population samples; these data were largely from around the year 2000. The relative order of control overweight/obesity prevalences across studies was not similar to that from data reported on the WHO Global database for BMI (Spearman's $\rho = 0.41$, $p = 0.08$).

that collected anthropometric information—around 40% of the data have not been presented before. Another advantage of pooling individual records is that it permits uniform categorization of data, as well as the assessment of the effects of potentially confounding factors. In this regard, adjustment for smoking and alcohol consumption did not greatly alter the risk estimates.

With respect to potential biases, participation rates were generally lower in controls than cases, and a particular concern is whether controls are representative of the populations from which cases were drawn. It is reassuring to note that pooling data from studies with control participation rates of 70% or more gave findings similar to those reported overall. Nonetheless, it is still possible that poor control participation could have influenced our findings since we cannot rule out the possibility that those with obesity-related health problems (*e.g.*, type 2 diabetes, cardiovascular disease, respiratory difficulties and chronic musculoskeletal problems) may have been (more or) less likely to participate. If the latter applied, the increased risks in the highest BMI category could be an artefact of differential case-control participation.

The rapidly changing prevalence of obesity is a growing public health problem, and to further investigate the issue, age-standardized data calculated from height and weight measurements were sourced from the World Health Organization Global Database on BMI (who.int/bmi/). Interestingly, the relative order of the overweight (25–29.99 kg m^{-2})/obesity ($\geq$30 kg m^{-2}) prevalence across studies among our controls and that of the corresponding country-specific WHO BMI prevalence from around the year 2000 are not strongly correlated (Spearman's $\rho = 0.41$, $p = 0.08$) (Fig. 4). WHO data place the USA, Germany and the UK at the top while our self-reported information rate the Czech Republic, USA and Italy as having the highest overweight/obesity prevalence. Whilst differences between our data and WHO are likely to be related to factors such as age, sex and time period, they serve to illustrate the rapidly changing patterns and wide variations that exist around the world.

Self-reports of anthropometric information is known to be inexact, with height tending to be overestimated and weight underestimated.[52] The nature of individual misreporting is likely to be complex, being related not only to their actual size but also to other factors such as age and sex. In a cohort of British adults, for instance, where self-reported and measured data were compared, height was overestimated most by older people, shorter men and heavier women, while the greatest underestimation of weight was amongst heavier men and women.[53] This tendency for people to report BMI closer to "normal" may have diluted our odds ratios. It is also possible that weight loss associated with lymphoma may have influenced the recall of cases differently to that of controls. Because of this, at interview subjects were either asked to recall their usual weight or their weight at a specified times before diagnosis/interview, and restricting the analyses to the 6 studies (NCI-SEER, Mayo Phase 1, British Columbia, UK, North Italy and Italy) where data were requested at 1 or 5 years prior to diagnosis yielded similar results to the findings overall.

Whilst BMI derived from height and weight acts as an easily obtained estimate of adiposity, its use as a marker of obesity has several potential weaknesses. Across different ethnic groups, for example, a given BMI may not correspond to the same proportions of body fat.[54] Moreover since the index was originally devised as a means of assessing average body composition among sedentary individuals of working age it may not truly reflect the degree of adiposity across the population as a whole. For instance, among the elderly where muscle mass may have started to decline, body fat mass may be underestimated by BMI whereas amongst athletes it may be overestimated. To account for the potential variation in BMI as a marker of body fat across different populations, we grouped our data according to study-specific control distributions

as well as WHO BMI categories. We also repeated the analyses restricting data to Caucasians, and to North American and Northern European studies combined. Sensitivity analyses were conducted too among persons aged 65 or less (71% of our subjects), and among those who were not regular heavy exercisers where this information was requested (NCI-SEER, British Columbia and HERPACC2). These additional investigations gave similar findings to the presented results. More specific estimates of adiposity may be derived from total body fat mass and, as a marker for abdominal fat distribution, waist-to-hip ratios, but such data were not obtained in the studies included here and have only rarely been investigated with respect to NHL elsewhere, showing little effect.[15,21,26]

In conclusion, this pooled analysis of case-control studies from 13 countries, crossing 3 continents, did not find an association between NHL and increased BMI. ORs were raised in studies from some countries, namely the US, Canada and Northern European nations, but even within this group, heterogeneity was observed, questioning the validity of a combined odds ratio. The findings presented here were based on individual data from a large number of subjects enrolled in 18 studies, pooling of which were accomplished through the InterLymph consortium. Some of the advantages of this pooled analysis include information on confounders and NHL subtypes as well as data on height and weight, the constituent components of BMI. One potential confounding factor not assessed here is diet but dietary data will be examined, in conjunction with BMI, in a future InterLymph pooled analysis. Such investigations may further elucidate whether NHL or its subtypes are associated with obesity *per se*.

References

1. Marti A, Marcos A, Martinez JA. Obesity and immune function relationships. Obes Rev 2001;2:131–40.
2. Samartin S, Chandra RK. Obesity, overnutrition and the immune system. Nutr Res 2001;21:243–62.
3. Holly EA, Lele C, Bracci PM, McGrath MS. Case-control study of non-Hodgkin's lymphoma among women and heterosexual men in the San Francisco Bay Area, California. Am J Epidemiol 1999; 150:375–89.
4. Wolk A, Gridley G, Svensson M, Nyren O, McLaughlin JK, Fraumeni JF, Adami HO. A prospective study of obesity and cancer risk (Sweden). Cancer Causes Control 2001;12:13–21.
5. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 2003;348:1625–38.
6. Pan SY, Johnson KC, Ugnat AM, Wen SW, Mao Y. Association of obesity and cancer risk in Canada. Am J Epidemiol 2004;159:259–68.
7. Bahl S, Cotterchio M, Kreiger N, Klar N. Antidepressant medication use and non-Hodgkin's lymphoma risk: no association. Am J Epidemiol 2004;160:566–75.
8. Cerhan JR, Bernstein L, Severson RK, Davis S, Colt JS, Blair A, Hartge P. Anthropometrics, physical activity, related medical conditions, and the risk of non-hodgkin lymphoma. Cancer Causes Control 2005;16:1203–14.
9. Oh SW, Yoon YS, Shin SA. Effects of excess weight on cancer incidences depending on cancer sites and histologic findings among men: Korea National Health Insurance Corporation Study. J Clin Oncol 2005;23:4742–54.
10. Rapp K, Schroeder J, Klenk J, Stoehr S, Ulmer H, Concin H, Diem G, Oberaigner W, Weiland SK. Obesity and incidence of cancer: a large cohort study of over 145,000 adults in Austria. Br J Cancer 2005;93:1062–7.
11. Willett EV, Skibola CF, Adamson P, Skibola DR, Morgan GJ, Smith MT, Roman E. Non-Hodgkin's lymphoma, obesity and energy homeostasis polymorphisms. Br J Cancer 2005;93:811–16.
12. Chiu BC, Gapstur SM, Greenland P, Wang R, Dyer A. Body mass index, abnormal glucose metabolism, and mortality from hematopoietic cancer. Cancer Epidemiol Biomarkers Prev 2006;15:2348–54.
13. Engeland A, Tretli S, Hansen S, Bjorge T. Height and body mass index and risk of lymphohematopoietic malignancies in two million Norwegian men and women. Am J Epidemiol 2007;165:44–52.
14. Chiu BC, Soni L, Gapstur SM, Fought AJ, Evens AM, Weisenburger DD. Obesity and risk of non-Hodgkin lymphoma (United States). Cancer Causes Control 2007;18:677–85.
15. Lim U, Morton LM, Subar AF, Baris D, Stolzenberg-Solomon R, Leitzmann M, Kipnis V, Mouw T, Carroll L, Schatzkin A, Hartge P. Alcohol, Smoking, and Body Size in Relation to Incident Hodgkin's and Non-Hodgkin's Lymphoma Risk. Am J Epidemiol 2007;166:697–708.
16. Paffenbarger RS, Jr, Wing AL, Hyde RT. Characteristics in youth predictive of adult-onset malignant lymphomas, melanomas, and leukemias: brief communication. J Natl Cancer Inst 1978;60:89–92.
17. Whittemore AS, Paffenbarger RS, Jr, Anderson K, Lee JE. Early precursors of site-specific cancers in college men and women. J Natl Cancer Inst 1985;74:43–51.
18. Franceschi S, Serraino D, Bidoli E, Talamini R, Tirelli U, Carbone A, La Vecchia C. The epidemiology of non-Hodgkin's lymphoma in the north-east of Italy: a hospital-based case-control study. Leuk Res 1989;13:465–72.
19. Moller H, Mellemgaard A, Lindvig K, Olsen JH. Obesity and cancer risk: a Danish record-linkage study. Eur J Cancer A 1994;30:344–50.
20. Zhang S, Hunter DJ, Rosner BA, Colditz GA, Fuchs CS, Speizer FE, Willett WC. Dietary fat and protein in relation to risk of non-Hodgkin's lymphoma among women. J Natl Cancer Inst 1999;91:1751–8.
21. Cerhan JR, Janney CA, Vachon CM, Habermann TM, Kay NE, Potter JD, Sellers TA, Folsom AR. Anthropometric characteristics, physical activity, and risk of non-Hodgkin's lymphoma subtypes and B-cell chronic lymphocytic leukemia: a prospective study. Am J Epidemiol 2002;156:527–35.
22. Samanic C, Gridley G, Chow WH, Lubin J, Hoover RN, Fraumeni JF, Jr. Obesity and cancer risk among white and black United States veterans. Cancer Causes Control 2004;15:35–43.
23. Bosetti C, Dal Maso L, Negri E, Talamini R, Montella M, Franceschi S, La Vecchia C. Re: body mass index and risk of malignant lymphoma in Scandinavian men and women. J Natl Cancer Inst 2005;97:860–1.
24. Chang ET, Hjalgrim H, Smedby KE, Akerman M, Tani E, Johnsen HE, Glimelius B, Adami HO, Melbye M. Body mass index and risk of malignant lymphoma in Scandinavian men and women. J Natl Cancer Inst 2005;97:210–8.
25. Lukanova A, Bjor O, Kaaks R, Lenner P, Lindahl B, Hallmans G, Stattin P. Body mass index and cancer: results from the Northern Sweden health and disease cohort. Int J Cancer 2006;118:458–66.
26. MacInnis RJ, English DR, Hopper JL, Giles GG. Body size and composition and the risk of lymphohematopoietic malignancies. J Natl Cancer Inst 2005;97:1154–7.
27. Fernberg P, Odenbro A, Bellocco R, Boffetta P, Pawitan Y, Adami J. Tobacco use, body mass index and the risk of malignant lymphomas—a nationwide cohort study in Sweden. Int J Cancer 2006;118:2298–302.
28. Samanic C, Chow WH, Gridley G, Jarvholm B, Fraumeni JF, Jr. Relation of body mass index to cancer risk in 362,552 Swedish men. Cancer Causes Control 2006;17:901–9.
29. Willett EV, Smith AG, Dovey GJ, Morgan GJ, Parker J, Roman E. Tobacco and alcohol consumption and the risk of non-Hodgkin lymphoma. Cancer Causes Control 2004;15:771–80.
30. Besson H, Brennan P, Becker N, De Sanjose S, Nieters A, Font R, Maynadie M, Foretova L, Cocco PL, Staines A, Vornanen M, Boffetta P. Tobacco smoking, alcohol drinking and Hodgkin's lymphoma: a European multi-centre case-control study (EPILYMPH). Br J Cancer 2006;95:378–84.
31. Becker N, Deeg E, Nieters A. Population-based study of lymphoma in Germany: rationale, study design and first results. Leuk Res 2004;28:713–24.
32. de Sanjose S, Shah KV, Domingo-Domenech E, Engels EA, Fernandez dS, Alvaro T, Garcia-Villanueva M, Romagosa V, Gonzalez-Barca E, Viscidi RP. Lack of serological evidence for an association between simian virus 40 and lymphoma. Int J Cancer 2003;104: 522–4.
33. Talamini R, Montella M, Crovatto M, Dal Maso L, Crispo A, Negri E, Spina M, Pinto A, Carbone A, Franceschi S. Non-Hodgkin's lymphoma and hepatitis C virus: a case-control study from northern and southern Italy. Int J Cancer 2004;110:380–5.
34. Tajima K, Hirose K, Inoue M, Takezaki T, Hamajima N, Kuroishi T. A model of practical cancer prevention for out-patients visiting a hospital: the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC). Asian Pac J Cancer Prev 2000;1:35–47.
35. Suzuki T, Matsuo K, Ito H, Hirose K, Wakai K, Saito T, Sato S, Morishima Y, Nakamura S, Ueda R, Tajima K. A past history of gastric ulcers and *Helicobacter pylori* infection increase the risk of gastric malignant lymphoma. Carcinogenesis 2006;27:1391–7.
36. Spinelli JJ, Ng C, Weber JP, Connors JM, Gascoyne RD, Lai A, Brooks-Wilson A, Le ND, Berry B, Gallagher RP. Organochlorines and risk of non-Hodgkin lymphoma. Int J Cancer 121:2767–75.
37. Jaffe E, Harris N, Stein H, Vardiman J. Tumours of the Haemopoietic and Lymphoid Tissues. IARC Press, Lyon 2001.

38. Morton LM, Turner JJ, Cerhan JR, Linet MS, Treseler PA, Clarke CA, Jack A, Cozen W, Maynadie M, Spinelli JJ, Costantini AS, Rudiger T, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the International Lymphoma Epidemiology Consortium (InterLymph). Blood 2007;110:695–708.
39. WHO expert committee on physical status. Physical status: the use and interpretation of anthropometry: Report of a WHO Expert Committee. Geneva, Switzerland: World Health Organization, 1995. Report No.: 854.
40. Morton LM, Hartge P, Holford TR, Holly EA, Chiu BC, Vineis P, Stagnaro E, Willett EV, Franceschi S, La Vecchia C, Hughes AM, Cozen W, et al. Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (InterLymph). Cancer Epidemiol Biomarkers Prev 2005;14:925–33.
41. Rothman N, Skibola CF, Wang SS, Morgan G, Lan Q, Smith MT, Spinelli JJ, Willett E, de Sanjose S, Cocco P, Berndt SI, Brennan P, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. Lancet Oncol 2006;7:27–38.
42. Morton LM, Zheng T, Holford TR, Holly EA, Chiu BC, Costantini AS, Stagnaro E, Willett EV, Dal Maso L, Serraino D, Chang ET, Cozen W, et al. Alcohol consumption and risk of non-Hodgkin lymphoma: a pooled analysis. Lancet Oncol 2005;6:469–76.
43. Wang SS, Slager SL, Brennan P, Holly EA, De Sanjose S, Bernstein L, Boffetta P, Cerhan JR, Maynadie M, Spinelli JJ, Chiu BC, Cocco PL, et al. Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10 211 cases and 11 905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). Blood 2007;109:3479–88.
44. Kricker A, Armstrong BK, Hughes AM, Goumas C, Smedby KE, Zheng T, Spinelli JJ, De Sanjose S, Hartge P, Melbye M, Willett EV, Becker N, et al. Personal sun exposure and risk of non Hodgkin lymphoma: a pooled analysis from the Interlymph Consortium. Int J Cancer 2007;122:144–54.
45. Breslow NE, Day NE. Statistical methods in cancer research, vol. 1. The analysis of case-control studies. Lyon: International Agency for Research on Cancer, 1980.
46. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.
47. Colditz GA, Burdick E, Mosteller F. Heterogeneity in meta-analysis of data from epidemiologic studies: a commentary. Am J Epidemiol 1995;142:371–82.
48. Delgado-Rodriguez M. Glossary on meta-analysis. J Epidemiol Commun Health 2001;55:534–6.
49. Stata Statistical Software. Release 9 [computer program]. College Station. Texas: Stata Corporation, 2005.
50. Skibola CF, Holly EA, Forrest MS, Hubbard A, Bracci PM, Skibola DR, Hegedus C, Smith MT. Body mass index, leptin and leptin receptor polymorphisms, and non-hodgkin lymphoma. Cancer Epidemiol Biomarkers Prev 2004;13:779–86.
51. Larsson SC, Wolk A. Obesity and risk of non-Hodgkin's lymphoma: a meta-analysis. Int J Cancer 2007;121:1564–70.
52. Gorber SC, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. Obes Rev 2007;8:307–26.
53. Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. Public Health Nutr 2002;5:561–5.
54. WHO Consultation on Obesity. Obesity: preventing and managing the global epidemic: report of a WHO consultation. Geneva, Switzerland: World Health Organization, 2000. Report No.: 894.

International Lung Cancer Consortium: Pooled Analysis of Sequence Variants in DNA Repair and Cell Cycle Pathways

Rayjean J. Hung,[1,2] David C. Christiani,[4] Angela Risch,[5] Odilia Popanda,[5] Aage Haugen,[6] Shan Zienolddiny,[6] Simone Benhamou,[7,8] Christine Bouchardy,[9] Qing Lan,[10] Margaret R. Spitz,[11] H.-Erich Wichmann,[12,13] Loic LeMarchand,[14] Paolo Vineis,[15] Giuseppe Matullo,[16] Chikako Kiyohara,[17] Zuo-Feng Zhang,[18] Benhnaz Pezeshki,[18] Curtis Harris,[10] Leah Mechanic,[10] Adeline Seow,[19] Daniel P.K. Ng,[19] Neonila Szeszenia-Dabrowska,[20] David Zaridze,[21] Jolanta Lissowska,[22] Peter Rudnai,[23] Eleonora Fabianova,[24] Dana Mates,[25] Lenka Foretova,[26] Vladimir Janout,[27] Vladimir Bencko,[28] Neil Caporaso,[10] Chu Chen,[29] Eric J. Duell,[1] Gary Goodman,[29] John K. Field,[30] Richard S. Houlston,[31] Yun-Chul Hong,[32] Maria Teresa Landi,[10] Philip Lazarus,[33] Joshua Muscat,[33] John McLaughlin,[3] Ann G. Schwartz,[34] Hongbing Shen,[35] Isabelle Stucker,[36] Kazuo Tajima,[37] Keitaro Matsuo,[37] Michael Thun,[38] Ping Yang,[39] John Wiencke,[40] Angeline S. Andrew,[41] Stephanie Monnier,[1] Paolo Boffetta,[1] and Paul Brennan[1]

[1]IARC, Lyon, France; [2]Samuel Lunenfeld Research Institute of Mount Sinai Hospital; [3]Cancer Care Ontario, Toronto, Ontario, Canada; [4]Harvard University, Boston, Massachusetts; [5]German Cancer Research Center, Heidelberg, Germany; [6]National Institute of Occupational Health, Oslo, Norway; [7]INSERM U794, Fondation Jean Dausset CEPH, Paris, France; [8]CNRS FRE2939, Institut Gustave Roussy, Villejuif, France; [9]Cancer Registry, Geneva, Switzerland; [10]National Cancer Institute, Rockville; [11]The University of Texas M. D. Anderson Cancer Center, Houston, Texas; [12]GSF-National Research Center for Environment and Health, Neuherberg, Germany; [13]IBE, University of Munich, Germany; [14]University of Hawaii, Hawaii, Honolulu; [15]Imperial College, London, United Kingdom; [16]University of Turin, Turin, Italy; [17]Kyushu University, Fukuoka, Japan; [18]University of California at Los Angeles, Los Angeles, California; [19]National University of Singapore, Singapore, Singapore; [20]Department of Epidemiology, Institute of Occupational Medicine, Lodz, Poland; [21]Institute of Carcinogenesis, Cancer Research Centre, Moscow, Russia; [22]Department of Cancer Epidemiology and Prevention, The M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland; [23]National Institute of Environmental Health, Fodor József National Center for Public Health, Budapest, Hungary; [24]Regional Authority of Public Health, Banská Bystrica, Slovakia; [25]Institute of Public Health, Bucharest, Romania; [26]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic; [27]Department of Preventive Medicine, Faculty of Medicine, Palacky University, Olomouc, Czech Republic; [28]Charles University of Prague, First Faculty of Medicine, Institute of Hygiene and Epidemiology, Prague, Czech Republic; [29]Fred Hutchinson Cancer Research Center, Seattle, Washington; [30]Roy Castle Lung Cancer Research Programme, Cancer Research Centre, The University of Liverpool, Liverpool, United Kingdom; [31]Section of Cancer Genetics, Institute of Cancer Research, Sutton, United Kingdom; [32]Seoul National University College of Medicine, Seoul, Korea; [33]Penn State College of Medicine, Hershey, Pennsylvania; [34]Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit, Michigan; [35]Nanjing Medical University School of Public Health, Nanjing, People's Republic of China; [36]INSERM U754 and UMR-S754, IFR69, Villejuif, France; [37]Aichi Cancer Center Research Institute, Aichi, Japan; [38]American Cancer Society, Atlanta, Georgia; [39]Mayo Clinic Cancer Center, Rochester, Minnesota; [40]University of California at San Francisco, San Francisco, California; and [41]Dartmouth Medical School, Lebanon, New Hampshire

Abstract

Background: **The International Lung Cancer Consortium was established in 2004. To clarify the role of DNA repair genes in lung cancer susceptibility, we conducted a pooled analysis of genetic variants in DNA repair pathways, whose associations have been investigated by at least 3 individual studies.**

Methods: **Data from 14 studies were pooled for 18 sequence variants in 12 DNA repair genes, including *APEX1, OGG1, XRCC1, XRCC2, XRCC3, ERCC1, XPD, XPF, XPG, XPA, MGMT*, and *TP53*. The total number of subjects included in the analysis for each variant ranged from 2,073 to 13,955 subjects.**

Results: **Four of the variants were found to be weakly associated with lung cancer risk with borderline significance: these were *XRCC3* T241M [heterozygote odds ratio (OR), 0.89; 95% confidence interval (95% CI), 0.79-0.99 and homozygote OR, 0.84; 95% CI, 0.71-1.00] based on 3,467 cases and 5,021 controls from 8 studies, *XPD* K751Q (heterozygote OR, 0.99; 95% CI, 0.89-1.10 and homozygote OR, 1.19; 95% CI, 1.02-1.39) based on 6,463 cases and 6,603 controls from 9 studies, and *TP53* R72P (heterozygote OR, 1.14; 95% CI, 1.00-1.29 and homozygote OR, 1.20; 95% CI, 1.02-1.42) based on 3,610 cases and 5,293 controls from 6 studies. *OGG1* S326C homozygote was suggested to be associated with lung cancer risk in Caucasians (homozygote OR, 1.34; 95% CI, 1.01-1.79) based on 2,569 cases and 4,178 controls from 4 studies but not in Asians. The other 14 variants did not exhibit main effects on lung cancer risk.**

Discussion: **In addition to data pooling, future priorities of International Lung Cancer Consortium include coordinated genotyping and multistage validation for ongoing genome-wide association studies.** (Cancer Epidemiol Biomarkers Prev 2008;17(11):3081–9)

Introduction

Lung cancer continues to be the most common cancer overall and the leading cause of cancer death worldwide. In 2006, there were an estimated 1,352,000 new cases and 1,179,000 deaths (1). Disease survival continues to be poor with a 5-year mortality of ~90%. The only current option for disease control is through avoidance of exposure to lung carcinogens. However, much research remains to be done among women, never smokers, and

the young (2, 3). This emphasizes the importance of further understanding its etiology, including carcinogenesis in subgroup of interests and among histopathologic subtypes.

With the aim of sharing comparable data from ongoing lung cancer case-control and cohort studies, as well as increasing power for focused analysis of special subgroups, we established the International Lung Cancer Consortium (ILCCO) in 2004. The overall objectives are to achieve greater power, especially for subgroup analyses, reduce duplication of research effort, replicate novel findings, and maximize the cost efficiency through large collaborative efforts.

The consortium is operated under the guidelines and the policies addressing issues of data sharing, intellectual properties, authorship, and organization. These guidelines and policies followed the general principles adopted by InterLymph (4) and were amended and approved by ILCCO members.[42] Working groups were formed to oversee research areas that were considered priorities for the consortium. The current working groups include (*a*) Genetic Susceptibility, (*b*) Family History, (*c*) Risks among Nonsmokers, (*d*) Rare Histologic Types, (*e*) Occupational Factors, (*f*) Medical Conditions, and (*g*) Statistical Analysis (a working group that provides consultation to other groups on a project-specific basis). The coordinators and current research projects of each working group are posted on the ILCCO Web portal.

As the first proof-of-principle study, we conducted pooled analysis on sequence variants in the pathways of DNA repair, a critical defense mechanism against human carcinogenesis, and cell cycle control. The DNA repair system maintains the integrity of the human genome by reducing the mutation frequency of cancer-related genes, minimizing replication errors, removing DNA damage, and minimizing deleterious rearrangements arising via aberrant recombination (5). Cells with damaged DNA must either pause in the cell cycle to allow for repair or succumb to elimination by apoptosis, and the activation of cell cycle checkpoints is a critical component of the cellular response to DNA damage (6). Defects in DNA repair and cell cycle control pathway are likely to play a crucial role in tobacco-related lung cancer *a priori*. It has been hypothesized that a combination of low-penetrance genetic variants may account for a proportion of the genetic component for lung cancer susceptibility, and the candidate gene approach has guided research in this field in the past decade. However, most of the associations in such studies have not been replicated. The likely reasons include lack of statistical power leading to false-negative and false-positive results due to multiple comparisons, the latter being exacerbated due to publication bias. To evaluate whether genetic variants in DNA repair and cell cycle control pathways might influence the predisposition to lung cancer, we studied 18 variants in 12 DNA repair enzymes in a total of 14 lung cancer studies in ILCCO.

Received 5/6/08; revised 7/28/08; accepted 8/12/08.
Grant support: National Cancer Institute R03 grant CA119704.
Requests for reprints: Paolo Boffetta, IARC, 150 cours Albert Thomas, 69008 Lyon, France. Phone: 33-47273-8441; Fax: 33-47273-8320. E-mail: boffetta@iarc.fr
Copyright © 2008 American Association for Cancer Research.
doi:10.1158/1055-9965.EPI-08-0411

Materials and Methods

Study Population. Investigators who had conducted epidemiologic studies of lung cancer were invited to participate in ILCCO and requested to complete a consortium questionnaire. The eligibility criteria of studies to be included in ILCCO were that they had a study protocol for subject recruitment and a structured questionnaire for lifestyle information. Lung cancer researchers with expertise in molecular biology, pathology, and other relevant fields were also invited to join the consortium. The consortium was established with funding from the National Cancer Institute (NCI) and the IARC.

Worldwide, 39 lung cancer studies have participated in ILCCO to date including 19 population-based case-control studies, 14 hospital-based case-control studies, 2 case-control studies with mixed types of controls, and 3 cohort studies, which comes up to a total of more than 46,000 case-control pairs. Fifteen studies were conducted in North America, 13 studies were conducted in Europe, and 11 studies were conducted in Asia and Oceania. The basic characteristics of these 39 studies are summarized at the Table 2.

Pooled Analysis of DNA Repair and Cell Cycle Control Pathways. Fourteen studies that participated in ILCCO had genetic data on DNA repair and cell cycle control pathways and contributed data to this pooled analysis as indicated in Table 1. Six studies were conducted in European countries, four in the United States, and four in Asia or the Pacific islands. Seven studies recruited hospital-based controls, whereas the other seven recruited population-based controls with one nested in a cohort. The control groups in most of the studies were frequency matched with cases on age and sex, whereas some also matched on ethnicity or residence area, and three studies did not apply any matching factors. Written informed consents were obtained from all study subjects, and the investigations were approved by the ethical review board at each study center.

Statistical Methods. The individual-level data from each participating study were sent to the IARC for data pooling. Data elements submitted from the studies included demographic variables (e.g., age, sex, ethnicity, and country of residence), tobacco exposures, family history of cancer, and histology classification of the cases. The data submitted from all 14 studies were checked for inadmissible values, aberrant distributions, inconsistencies, and missing values. Queries were sent to the investigators to resolve all discrepancies and possible errors. Subjects with unknown age or sex were excluded from the analysis. Thus, data from 14 studies with a total of 8,454 lung cancer cases and 9,344 controls were pooled for the present project.

The frequency distribution of demographic variables and putative risk factors of lung cancer, including age, sex, ethnicity, and smoking, was examined for cases and controls. The ethnicity of the subjects were categorized according to the NIH definition as non-Hispanic Whites, Blacks or African Americans, Hispanic or Latinos,

[42] http://ilcco.iarc.fr

Table 1. Summary of studies participating in ILCCO

	Study	Principal investigator	Control source	Study period	Study location	Cases, *n*	Controls/ cohort, *n*
North America	Family Health Study-NS	A.G. Schwartz	Population	1984-1987	Detroit, MI	305	308
	Family Health Study-Young	A.G. Schwartz	Population	1990-2003	Detroit, MI	702	871
	Mayo-P	P. Yang	Population	1997-2006	United States	400	2,000
	University of California at San Francisco	J. Wiencke	Population	1998-2002	San Francisco, CA	1,011	900
	University of California at Los Angeles (29)*	Z.F. Zhang	Population	1999-2004	Los Angeles, CA	600	600
	Women's Epidemiology of Lung Disease	A.G. Schwartz	Population	2001-2005	Detroit, MI	576	575
	New England Lung Cancer Study	E. Duell	Population	2005-2008	New Hampshire, USA	285	250
	NCI-Maryland (30)*	C. Harris	Mixed	1998-2004	Baltimore, MD	800	800
	Samuel Lunenfeld Research Institute	J. McLaughlin	Mixed	1997-2002	Toronto, Ontario, Canada	516	660
	M. D. Anderson (24)*	M. Spitz	Hospital	1992-2004	Houston, TX	3,000	3,000
	Harvard (31)*	D. Christiani	Hospital	1992-2008	Boston, MA	3,300	3,300
	Mayo-H	P. Yang	Hospital	1997-2006	United States	4,000	2,000
	Moffitt	P. Lazarus	Hospital	1999-2003	Florida, USA	500	910
	Cancer Prevention Study-II	M. Thun	Cohort	1982-	United States	150	184,000
	CARET	G. Goodman	Cohort	1985-	United States	1,275	18,314
Europe	Norway (14)*	A. Haugen	Population	1986-	Norway	2,000	2,000
	GenAir (10) (European Prospective Investigation into Cancer and Nutrition)*	P. Vineis	Population	1993-1998	10 European countries	271	2,977
	Danish Cancer Society	H. Skuladottir	Population	1998-2001	Denmark	281	281
	Liverpool Lung Project	J. Field	Population	1998-2016	Liverpool, UK	1,500	3,000
	Lung Cancer of Young (32)*	H.E. Wichmann	Population	2000-2004	Germany	800	2,000
	deCODE	T. Rafnar	Population	2000-2012	Iceland	800	3,000
	INSERM-P	I. Stucker	Population	2002-2005	France	3,000	3,000
	NCI-Italy	M.T. Landi	Population	2002-2005	Lombardy, Italy	2,101	2,120
	INSERM-CEPH (11)*	S. Benhamou	Hospital	1987-1992	Paris, France	151	171
	Lung Cancer Study	I. Stucker	Hospital	1990-1992	Paris, France	310	301
	German Cancer Research Center (33)*	A. Risch	Hospital	1996-	Heidelberg, Germany	1,200	650
	IARC (12)*	P. Boffetta	Hospital	1998-2002	Central/eastern Europe	2,861	3,118
	Institute of Cancer Research	R. Houlston	Hospital		Sutton, UK	6,000	10,000
Asia and Oceania	NCI-China 1	Q. Lan	Population	1985-1990	Xuan Wei, People's Republic of China	500	500
	Hawaii-P (13)*	L. Le Marchand	Population	1992-1997	Hawaii, USA	582	582
	Kyushu (34)*	C. Kiyohara	Population	1994-1996	Japan	192	130
	NCI-China 2 (35)*	Q. Lan	Population	1995-1996	Xuan Wei, People's Republic of China	122	122
	Singapore 1 (36)*	A. Seow	Hospital	1996-1998	Singapore	303	765
	Aichi	K. Tajima/K. Mastuo	Hospital	2000-2003	Aichi, Japan	400	800
	Seoul	Y.C. Hong	Hospital	2001-2008	Seoul, Korea	1,150	1,150
	Eastern China	H. Shen	Hospital	2002-2007	Eastern China	2,000	2,000
	Singapore 2	A. Seow	Hospital	2005-2007	Singapore	400	800
	Hawaii	L. Le Marchand	Cross-sectional	2000-2004	Hawaii, USA	600	
	NCI-China cohort	Q. Lan	Cohort	1975-1995	People's Republic of China	1,600	42,663
Total						46,544	300,618

*Indicates the 14 studies contributing to the DNA repair pooled analysis with references provided.

Asians, Native Hawaiians or other Pacific islanders, American Indians, or others. Former smokers were defined as smokers who quit smoking at least 2 years before interview or diagnosis when the exact duration of time since quitting was available or based on self-reports when the duration of quitting was not available. Cumulative tobacco consumption was calculated as the product of smoking duration and intensity and expressed as pack-years.

Data from 14 studies were pooled for 18 sequence variants in 12 DNA repair genes, including *APEX1* D148E; *OGG1* S326C; *XRCC1* R194W, R280H, and R399Q; *XRCC2* R188H; *XRCC3* T241M; *ERCC1* T354C and C8092A; *XPD* D312N and K751Q; *XPF* R415Q; *XPG* H11054D; *XPA* G23A; *MGMT* L84F, I143V, and K178R; and *TP53* R72P. We assessed the deviations from Hardy-Weinberg equilibrium in the control population for individual studies. Studies in which the allele frequency among control group departed from Hardy-Weinberg equilibrium with $P < 0.01$ were excluded from the analysis, including the Norway study from the analysis of *APEX1* D148E, *OGG1* S326C, and *XPD* K751Q (Hardy-Weinberg equilibrium $P < 0.0001$ for all markers mentioned above) and the Harvard study from the analysis of *OGG1* S326C (Hardy-Weinberg equilibrium $P = 0.006$).

Genotypes were categorized into three groups (major allele homozygous, heterozygous, and homozygous variant). We estimated the genotype-specific odds ratios (OR), OR per allele, and their associated confidence intervals (95% CI) of lung cancer in each study using unconditional logistic regression modeling, adjusting for age, sex, cumulative tobacco smoking (expressed as pack-years), and country (when the study was conducted in multiple countries). When information on cumulative tobacco smoking was missing, it was imputed using the median of the study-specific control population. When there were at least three studies available, the summary estimates were obtained using a two-stage random-effect model, which allows for unexplained sources of heterogeneity among studies (7). Studies in which the OR could not be estimated (because one or more cells in the 4-fold table had no subjects) were excluded from the pooled analysis. The number of studies may appear to be different in each stratum depending on the amount of data provided in each stratum. A test of heterogeneity based on Q statistics was done for each summary estimate.

When there was evidence of heterogeneity across the study-specific ORs, we evaluated the source of heterogeneity by meta-regression and by stratified analysis on ethnicity and type of controls (8). If the heterogeneity was not due to any study characteristic, we conducted influence analysis and evaluated the source of heterogeneity from any single study by a Galbraith plot and comparing the Q values. The study contributing the most heterogeneity was excluded from the summary estimate.

We conducted stratified analyses by histology of lung cancer to investigate the potential modification in the effect of each polymorphism by histologic subtype. We also evaluated the modulating effects of tobacco smoking and family history of cancer by stratifying and comparing the strata-specific estimates. All statistical analysis above was conducted with STATA software version 9.

For genes that have multiple single nucleotide polymorphisms that are in linkage disequilibrium ($D' > 0.7$) such as *XRCC1* (R194W, R280H, and R399Q), *ERCC1* (T354C and C8092A), *XPD* (D312N and K751Q), and *MGMT* (I143V and K178R), we have also conducted haplotype analysis based on the pooled data set. Haplotype dosages for variants in linkage disequilibrium were estimated based on E-M algorithm by the tag single nucleotide polymorphism program to indicate an individual's probability of being heterozygote or homozygote and for a log-additive model (9). The haplotype dosage was then analyzed as a continuous variable in multivariate logistic regression adjusting for age, gender, geographic area, and study.

Results

The demographic distribution of the pooled data set for DNA repair genes is shown in Table 2. Individual studies contributed between 1.4% and 26.6% of cases and similar range of controls. More than 60% of the cases and controls were males, and the majority of the subjects were ages >60 years. More than 85% of the cases and controls were non-Hispanic Whites, and ~6% to 7% of the cases and controls were Asians. As expected, the prevalence of smoking was higher among cases than controls. Squamous cell carcinoma and adenocarcinoma were the two predominant histologic subtypes of lung cancer in this analysis.

Table 3 shows the summary ORs for each variant. The total number of subjects included in the analysis for each variant ranged from 2,073 to 13,955. Overall, 2 of the 18 variants were suggested to be associated with lung cancer risk: *XRCC3* T241M (heterozygote OR, 0.89; 95% CI, 0.79-0.99 and homozygote OR, 0.84; 95% CI, 0.71-1.00; P value per allele = 0.01) based on 3,467 cases and 5,021 controls from 8 studies and *TP53* R72P (heterozygote OR, 1.14; 95% CI, 1.00-1.29 and homozygote OR, 1.20; 95% CI, 1.02-1.42; P value per allele = 0.01) based on 3,610 cases and 5,293 controls from 6 studies. The study-specific ORs for *XRCC3* T241M and *TP53* R72P are shown in Fig. 1. In addition, we observed a weak association for two other variants but among homozygote carriers only: *XPD* K751Q and *OGG1* S326C. *XPD* K751Q homozygote carriers conferred an OR (95% CI) of 1.19 (1.02-1.39) based on 6,463 cases and 6,603 controls from 9 studies. There was evidence of heterogeneity for *OGG1* S326C homozygote ($P = 0.02$), which was no longer present after stratification on ethnicity. *OGG1* 326C/326C genotype was shown to be associated with lung cancer risk in non-Hispanic Whites (homozygote OR, 1.34; 95% CI, 1.01-1.79) based on 2,569 cases and 4,178 controls from 4 studies (10-13) but not in Asians. The other 14 variants did not appear to have main effects on lung cancer risk.

Heterogeneity among studies was observed for *XRCC2* R188H (heterogeneity P value for the 188R/188H < 0.0001) and *ERCC1* T354C (heterogeneity $P = 0.002$ for 354T/354C and 0.001 for 354C/354C). This heterogeneity was not explained by either ethnicity, control source, or genotyping methodology but was accounted for by a single outlying study (14). Exclusion of this study did not change the conclusion of the summary estimates but reduced the heterogeneity to the P values of 0.52, 0.42, and 0.80 for *XRCC2* R188H

Table 2. Frequency distribution of demographic variables and putative risk factors of lung cancer in ILCCO pooled data set for DNA repair pooled analysis

	Cases, *n* (%)	Controls, *n* (%)
Total	8,454	9,344
Study ID name (genotyping method)		
Norway (TaqMan and APEX; ref. 14)	334 (3.95)	413 (4.4)
GenAir (TaqMan and PE/DHPLC; ref. 10)	116 (1.4)	1,077 (11.5)
Kyushu (RFLP; ref. 34)	190 (2.3)	108 (1.2)
University of California at Los Angeles (RFLP; ref. 29)	497 (5.9)	902 (9.7)
Lung Cancer of Young (MassArray-PE; ref. 32)	270 (3.2)	222 (2.4)
NCI-Maryland (TaqMan; ref. 30)	490 (5.8)	678 (7.3)
INSERM-CEPH (BeadArray; ref. 11)	151 (1.8)	172 (1.8)
M. D. Anderson (TaqMan; ref. 24)	342 (4.0)	362 (3.9)
Harvard (TaqMan; ref. 31)	2,253 (26.6)	1,418 (15.2)
German Cancer Research Center (LightCycler and RFLP; ref. 33)	1,020 (12.1)	425 (4.6)
IARC (TaqMan and Amplifluor; ref. 12)	2,210 (26.1)	2,845 (30.5)
NCI-China 2 (TaqMan; ref. 35)	119 (1.4)	114 (1.2)
Singapore 1 (TaqMan; ref. 36)	125 (1.5)	162 (1.7)
Hawaii-P (RFLP; ref. 13)	337 (4.0)	446 (4.8)
Sex		
Male	5,332 (63.1)	5,698 (61.0)
Female	3,122 (36.9)	3,646 (39.0)
Age (y)		
<40	163 (1.9)	431 (4.6)
40-49	1,152 (13.6)	1,539 (16.5)
50-59	2,362 (27.9)	2,838 (30.4)
60-69	2,693 (31.8)	2,854 (30.5)
70-79	1,835 (21.7)	1,568 (16.8)
≥80	249 (2.9)	114 (1.2)
Smoking status		
Never	878 (10.4)	3,326 (35.6)
Ever	7,547 (89.3)	5,999 (64.2)
Former	2,688 (31.8)	3,211 (34.4)
Current	4,788 (56.6)	2,721 (29.1)
Education		
Low	1,249 (18.7)	1,715 (21.0)
Medium	3,421 (51.2)	3,761 (46.0)
High	2,006 (30.0)	2,698 (33.0)
Ethnicity		
Non-Hispanic Whites	7,404 (87.6)	8,033 (86.0)
Asians	600 (7.1)	605 (6.5)
Native Hawaiian/other Pacific Islander	85 (1.0)	102 (1.1)
Black or African American	251 (3.0)	359 (3.8)
Hispanic or Latino	96 (1.1)	213 (2.3)
American Indian or Alaska Native	5 (0.1)	3 (0.03)
Other	8 (0.1)	27 (0.3)
Histology		
Squamous cell carcinoma	2,399 (28.4)	
Small cell carcinoma	1,052 (12.4)	
Large cell carcinoma	498 (5.9)	
Adenocarcinoma	2,923 (34.6)	
Others, not otherwise specified	1,416 (12.9)	

NOTE: Education level: low, no education to junior high school; medium, high school or technical school level; high, university level and above. PE, primer extension; APEX, arrayed primer extension.

heterozygote and *ERCC1* T354C heterozygote and homozygote, respectively.

Table 4 shows the stratified estimates of *XRCC3* T241M and *TP53* R72P by ethnicity, smoking, histology, and gender. There were no differences in the stratum-specific estimates when stratified by ethnicity and gender. However, both variants appeared to confer a stronger association with lung cancer risk among smokers. In terms of the effect on the histologic subtypes, *XRCC3* T241M allele showed a more prominent effect on risk of small cell carcinoma (heterozygote OR, 0.84; 95% CI, 0.66-1.08 and homozygote OR, 0.66; 95% CI, 0.44-0.98), whereas *TP53* 72P allele showed a more prominent effect on risk of squamous cell carcinoma (heterozygote OR, 1.26; 95% CI, 1.04-1.52 and homozygote OR, 1.52; 95% CI, 1.19-1.94).

Haplotype analysis suggested that subjects who carried the *XPD* 312N-751Q haplotype had an increased risk of lung cancer with an OR (95% CI) of 1.19 (1.03-1.37) when the subjects carried two copies of such haplotype. We did not observe any haplotype-specific association for *APEX1*, *XRCC1*, and *ERCC1* (data not shown).

Discussion

We have established the foundation of international collaboration in the area of molecular and genetic epidemiology of lung cancer. Here, we pooled the genotype data for 18 sequence variants that are commonly investigated for lung cancer from studies in the United

Table 3. Summary estimates of the main effects of the genetic variants

Gene variant (dbSNP no.)	No. studies (reference)	Case no. GG/Gg/gg	Control no. GG/Gg/gg	Heterozygotes OR (95% CI)	Homozygotes OR (95% CI)	Per allele OR (95% CI)	P_{trend}	$P_{heterogeneity}$
APEX1 D148E (rs3136820)	6 (10-12, 31, 33, 35)	1,499/2,480/1,101	1,596/2,903/1,361	0.89 (0.81-0.99)	0.91 (0.78-1.06)	0.95 (0.88-1.02)	0.14	0.34
OGG1 S326C (rs1052133)*	4 (10-13)	1,648/805/116	2,628/1,389/161	0.95 (0.84-1.07)	1.34 (1.01-1.79)	1.02 (0.93-1.13)	0.63	0.82
XRCC1 R194W (rs1799782)	5 (10-12, 14, 35)	2,495/363/26	3,947/596/28	0.92 (0.79-1.08)	1.57 (0.76-3.26)	0.98 (0.85-1.13)	0.77	0.54
XRCC1 R280H (rs25489)	4 (12, 14, 32, 35)	2,568/273/16	3,123/351/9	1.01 (0.75-1.36)	2.06 (0.83-5.09)	1.06 (0.83-1.37)	0.62	0.21
XRCC1 R399Q (rs25487)	9 (10-12, 14, 31-35)	2,768/2,871/793	2,824/2,934/792	0.98 (0.88-1.08)	0.93 (0.75-1.14)	0.97 (0.89-1.05)	0.39	0.26
XRCC2 R188H (rs3218536)†	3 (10-12)	2,126/281/10	3,324/470/18	1.01 (0.84-1.21)	0.90 (0.35-2.28)	0.99 (0.84-1.17)	0.90	0.74
XRCC3 T241M (rs861539)	8 (10-12, 14, 33-36)	1,719/1,386/362	2,144/2,248/629	**0.89 (0.79-0.99)**	**0.84 (0.71-1.00)**	**0.91 (0.84-0.98)**	**0.01**	0.50
ERCC1 T354C (rs11615)†	3 (10, 12, 31)	1,730/2,052/684	2,034/2,373/811	1.04 (0.94-1.15)	1.03 (0.89-1.19)	1.02 (0.95-1.09)	0.56	0.64
ERCC1 C8092A (rs3212986)	3 (12, 14, 31)	2,660/1,719/309	2,569/1,694/288	1.00 (0.91-1.11)	1.06 (0.88-1.29)	1.02 (0.94-1.09)	0.66	0.61
XPD D312N (rs1799793)	9 (10, 12, 14, 24, 30-33, 35)	2,875/2,988/1,009	2,948/3,124/1,011	0.96 (0.86-1.06)	1.09 (0.97-1.23)	1.03 (0.97-1.08)	0.37	0.83
XPD K751Q (rs13181)	9 (10-12, 24, 31, 33-36)	2,592/2,869/1,002	2,650/2,995/958	0.99 (0.89-1.10)	1.19 (1.02-1.39)	1.07 (0.98-1.18)	0.15	0.05
XPF R415Q (rs1800067)	3 (11, 12, 14)	2,201/306/13	2,208/390/21	1.00 (0.84-1.19)	0.72 (0.24-2.13)	0.96 (0.82-1.13)	0.65	0.66
XPG H1104D (rs17655)	5 (11, 12, 29, 24, 35)	1,852/1,155/209	2,485/1,510/286	1.07 (0.91-1.25)	1.00 (0.81-1.25)	1.03 (0.93-1.13)	0.62	0.32
XPA G23A (rs1800975)	3 (12, 14, 33)	1,181/1,260/362	1,422/1,603/427	0.85 (0.66-1.10)	1.06 (0.82-1.37)	0.97 (0.82-1.15)	0.70	0.08
MGMT L84F (rs12917)	3 (10, 14, 35)	366/143/19	1,121/391/33	0.92 (0.60-1.41)	1.53 (0.76-3.08)	1.02 (0.73-1.42)	0.91	0.13
MGMT I143V (rs2308321)	3 (11, 12, 14)	1,696/538/45	2,594/675/51	1.06 (0.94-1.20)	1.05 (0.91-1.20)	1.19 (0.71-1.99)	0.50	0.31
MGMT K178R (rs2308327)	3 (12, 14, 35)	1,800/528/108	2,342/621/112	1.08 (0.93-1.26)	1.27 (0.63-2.54)	1.07 (0.91-1.25)	0.41	0.31
TP53 R72P (rs1042522)	6 (10, 12, 13, 30, 33, 36)	1,764/1,480/366	2,733/2,079/481	**1.14 (1.00-1.29)**	**1.20 (1.02-1.42)**	**1.10 (1.02-1.18)**	**0.01**	0.40

NOTE: OR, individual OR adjusted for age, sex, cumulative tobacco smoking (as pack-years), and country when applicable. GG/Gg/gg, major allele homozygotes/heterozygotes/minor allele homozygotes.
*Estimates in non-Hispanic Whites only due to the significant heterogeneity by ethnicity and limited data on Asians.
†Norway study was excluded as it conferred an outlying estimate and contributed to the majority of the heterogeneity.

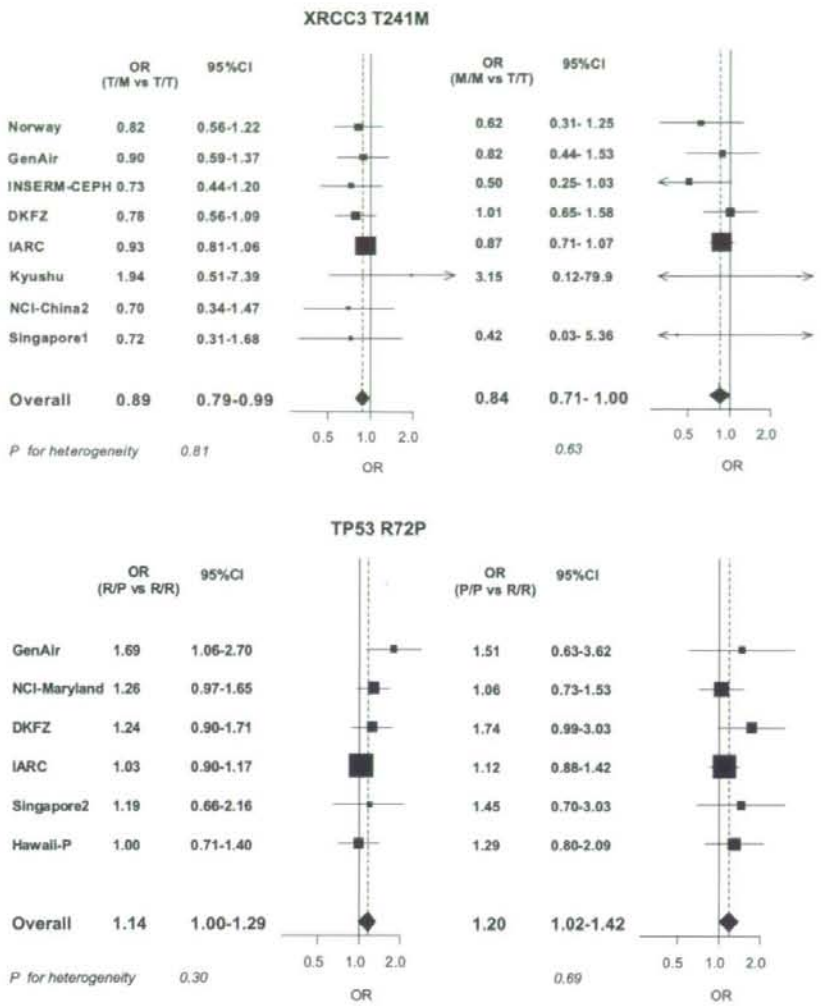

Figure 1. Study-specific estimates and summary OR for *XRCC3* T241M and TP53 R72P.

States, Europe, and Asia. None of the variants appeared to have a large effect on lung cancer risk, although we did observe a modest association of *XRCC3* T241M and *TP53* R72P polymorphisms and lung cancer risk overall as well as the effect of *XPD* haplotype 312N-751Q and *OGG1* 326C/326C genotype on lung cancer risk among non-Hispanic Whites. The potential associations between lung cancer risk and the other 14 variants were refuted based on this large analysis.

TP53 gene is one of the most studied human genes due to its critical role as tumor suppressor gene, and we observed an increased risk of lung cancer among TP53 72Pro allele carriers. The 72Pro allele has been suggested to be less efficient in suppressing cell transformation and to induce apoptosis with slower kinetics (15, 16). Several studies have investigated the association between this variant and lung cancer risk; however, the results have been inconsistent as reviewed by Matakidou et al. (17). We observed a more apparent effect of 72Pro allele among smokers, which agrees with the stronger association with the risk of squamous cell carcinoma. These results suggest that the effect of TP53 72Pro allele on lung cancer risk is mainly present in an environment challenged by tobacco-related carcinogens. It has also been shown previously that TP53 somatic mutations increase the risk of lung squamous cell carcinoma when compared with adenocarcinoma (18).

We have observed a modest protective effect conferred by XRCC3 241Met allele carriers. Again, the association is mainly present among smokers and tobacco-related histology, such as small cell carcinoma and squamous cell carcinoma with no evidence of heterogeneity across ethnicities. XRCC3 is a protein of Rad51-related family, which participates in homologous recombination repair of the double-strand breaks (19). Previous studies showed that the 241Met allele was shown to increase

Table 4. Stratified analysis for XRCC3 T241M and TP53 R72P by ethnicity, smoking status, and histology

Gene variant	No. studies	No. cases	No. controls	Heterozygotes OR (95% CI)	Homozygotes OR (95% CI)	Per allele OR (95% CI)	P_{trend}	$P_{heterogeneity}$
XRCC3 T241M								
Ethnicity								
Non-Hispanic Whites*	5	3,042	4,644	0.89 (0.80-0.99)	0.84 (0.71-1.00)	0.91 (0.84-0.98)	0.01	0.54
Asian*	3	425	377	0.82 (0.49-1.37)	0.92 (0.12-6.98)	0.84 (0.48-1.49)	0.55	0.22
Smoking								
Never†	6	477	1,951	1.03 (0.78-1.36)	1.28 (0.86-1.92)	1.10 (0.91-1.33)	0.32	0.50
Ever*	7	2,985	3,062	0.85 (0.76-0.96)	0.78 (0.65-0.94)	0.87 (0.80-0.95)	0.002	0.73
Former*	5	686	1,410	0.79 (0.63-0.98)	0.77 (0.55-1.09)	0.86 (0.73-1.00)	0.05	0.93
Current*	5	2,229	1,587	0.91 (0.78-1.05)	0.80 (0.54-1.20)	0.90 (0.79-1.03)	0.11	0.32
Histology								
Squamous cell*	5	1,350	3,832	0.88 (0.71-1.11)	0.78 (0.62-0.99)	0.90 (0.74-1.08)	0.24	0.14
Small cell*	4	414	3,426	0.84 (0.66-1.08)	0.66 (0.44-0.98)	0.83 (0.70-0.98)	0.03	0.82
Adenocarcinoma*	6	935	4,741	0.88 (0.74-1.05)	1.04 (0.80-1.34)	0.98 (0.87-1.10)	0.72	0.87
Gender								
Male‡	7	2,546	3,348	0.87 (0.76-0.99)	0.83 (0.69-1.01)	0.90 (0.82-0.98)	0.02	0.49
Female‡	7	851	1,664	0.91 (0.74-1.13)	0.90 (0.64-1.26)	0.94 (0.81-1.10)	0.44	0.44
TP53 R72P								
Ethnicity								
Non-Hispanic Whites*	5	3,159	4,605	1.14 (0.94-1.39)	1.19 (0.97-1.45)	1.12 (0.99-1.26)	0.07	0.20
Asian*	2	232	332	1.30 (0.87-1.95)	1.46 (0.87-2.47)	1.23 (0.95-1.59)	0.92	0.85
Smoking								
Never†	6	424	2,082	1.16 (0.91-1.48)	0.96 (0.64-1.43)	1.03 (0.87-1.22)	0.69	0.59
Ever*	5	3,181	3,203	1.17 (0.96-1.41)	1.37 (1.05-1.77)	1.18 (1.03-1.35)	0.017	0.11
Former*	5	929	1,804	1.23 (1.03-1.48)	1.37 (1.02-1.85)	1.19 (1.05-1.36)	0.007	0.42
Current*	4	2,251	1,397	1.11 (0.83-1.49)	1.42 (0.83-2.42)	1.17 (0.91-1.50)	0.20	0.03
Histology								
Squamous cell*	5	1,296	4,466	1.26 (1.04-1.52)	1.52 (1.19-1.94)	1.22 (1.10-1.35)	<0.001	0.59
Small cell*	3	376	3,635	0.81 (0.63-1.05)	1.27 (0.68-2.37)	0.96 (0.80-1.15)	0.64	0.41
Adenocarcinoma*	6	1,061	5,294	1.05 (0.90-1.22)	1.00 (0.77-1.29)	1.02 (0.91-1.14)	0.75	0.60
Gender								
Male‡	5	2,495	3,307	1.09 (0.96-1.23)	1.30 (1.05-1.60)	1.12 (1.02-1.22)	0.01	0.49
Female‡	6	1,115	1,987	1.14 (0.95-1.36)	1.10 (0.83-1.46)	1.08 (0.95-1.22)	0.25	0.75

*OR adjusted for age, sex, tobacco cumulative exposure (pack-years), and country when applicable.
†OR adjusted for age, sex, and country when applicable.
‡OR adjusted for age, tobacco cumulative exposure (pack-years) and country when applicable.

the DNA adduct level but had no effect on the repair of UV light -induced damage (20, 21). Two meta-analysis were conducted previously and reported a null association (21, 22). However, neither of them was able to adjust the results by a standard set of covariates across studies nor stratification by histology or smoking status, which may at least partially explain the differences in the conclusions from the present analysis.

The major strength of the pooled analysis in the consortium was to increase statistical power for common sequence variations with possibly modest effects, particularly for analysis in subgroups of interests such as rare histology, never smokers, or familial cases. In addition, pooling individual data has the advantage of being able to conduct analysis based on a standard approach as well as including multiple markers such as haplotype analysis.

There are several limitations of the present pooled analysis. First, because the studies were conducted in different populations and did not follow a standard protocol, the validity of the pooled estimate can be also threatened by the heterogeneity of the studies. In addition to single outlying estimates, heterogeneity can result from differences among study populations, study design, and often methodologic aspects including genotyping methodology. Apart from using random-effects models to allow for study heterogeneity to be taken into account, we also conducted stratified analysis by study design when there was evidence of heterogeneity. However, we did not observe different effects by control source or genotyping methodology (data not shown). Second, the pooled data set contains subjects with different ethnic ancestry, mainly European decedents and Asians, which might lead to bias from population stratification or simply population mixing and mask the true association. We have conducted stratified analysis by ethnicity whenever appropriate but did not observe any differential effect by ethnicity, except for *OGG1* S326C, for which we reported the results of non-Hispanic Whites only, as they contributed the majority of the data for this variant. Third, the pooled analysis is limited to existing data available in at least three of the studies in ILCCO. Therefore, we were not able to conduct comprehensive investigations using tagging single nucleotide polymorphisms of specific genes of interest. For example, another TP53 variant of 16-bp repeats located in intron 3, which is in linkage disequilibrium with the R72P allele, has been hypothesized to increase lung cancer risk (23, 24). However, we were not able to disentangle the effect conferred by these two linked variants in the present analysis due to lack of relevant data in the participating studies. This limitation can be overcome in the consortium by coordinated genotyping instead of simply data pooling, and this is currently under way in ILCCO.

Future Research Priorities and Consortium Values. ILCCO provides an opportunity for leading researchers

of lung cancer epidemiology to share results, plan pooled analyses, and discuss replication studies. This study shows the value of consortia for clarifying putative risk associated with complex diseases (25). Other ongoing research activities in ILCCO include pooled analysis of risk factors for rare histologic types of lung cancer, pooled analysis of data on family history, occupational exposures, and indoor air pollution.

Future prospects of ILCCO include multistage validation and fine mapping for possible causative genetic regions identified from ongoing genome-wide association studies (26-28). In this respect, consortia and international collaborations are developing into an ideal way to maximize study efficiency and overcome the limitations (particularly in terms of statistical power) of individual studies. We anticipate that ILCCO will be a major step toward improving our understanding of the causes and mechanisms of lung cancer and the beginning of a long-standing cooperation.

Acknowledgments

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Siti Nurbaya for the technical assistance to the Singapore study and all the investigators in the European Prospective Investigation into Cancer and Nutrition study for the contribution of the data.

References

1. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002: cancer incidence, mortality and prevalence worldwide. Lyon: IARC Press; 2004.
2. Peto R, Darby S, Deo H, et al. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. BMJ 2000;321:323–9.
3. Boffetta P, Trichopoulos D. Cancer of the lung, larynx, and pleura. In: Adami HO, Hunter D, Trichopoulos D, editors. Text book of cancer epidemiology. New York: Oxford University Press; 2002. p. 248–80.
4. Boffetta P, Armstrong B, Linet M, et al. Consortia in cancer epidemiology: lessons from InterLymph. Cancer Epidemiol Biomarkers Prev 2007;16:197–9.
5. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. Nature 2001;411:366–74.
6. Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. Nature 2004; 432:316–23.
7. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
8. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med 2002;21:589–624.
9. Carlson CS, Eberle MA, Rieder MJ, et al. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet 2004;74: 106–20.
10. Peluso M, Hainaut P, Airoldi L, et al. Methodology of laboratory measurements in prospective studies on gene-environment interactions: the experience of GenAir. Mutat Res 2005;574:92–104.
11. Feyler A, Voho A, Bouchardy C, et al. Point: myeloperoxidase -463G->A polymorphism and lung cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1550–4.
12. Hung RJ, Brennan P, Canzian F, et al. Large-scale investigation of base excision repair genetic polymorphisms and lung cancer risk in a multicenter study. J Natl Cancer Inst 2005;97:567–76.
13. Le Marchand L, Donlon T, Lum-Jones A, Seifried A, Wilkens LR. Association of the hOGG1 $Ser^{326}Cys$ polymorphism with lung cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:409–12.
14. Zienolddiny S, Campa D, Lind H, et al. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. Carcinogenesis 2006;27:260–7.
15. Dumont P, Leu JI, Della PA, III, George DL, Murphy M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet 2003;33:357–65.
16. Thomas M, Kalita A, Labrecque S, et al. Two polymorphic variants of wild-type p53 differ biochemically and biologically. Mol Cell Biol 1999;19:1092–100.
17. Matakidou A, Eisen T, Houlston RS. TP53 polymorphisms and lung cancer risk: a systematic review and meta-analysis. Mutagenesis 2003;18:377–85.
18. 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.
19. Liu Y, Tarsounas M, O'Regan P, West SC. Role of RAD51C and XRCC3 in genetic recombination and DNA repair. J Biol Chem 2007; 282:1973–9.
20. Matullo G, Palli D, Peluso M, et al. XRCC1, XRCC3, XPD gene polymorphisms, smoking and ^{32}P DNA adducts in a sample of healthy subjects. Carcinogenesis 2001;22:1437–45.
21. Manuguerra M, Saletta F, Karagas MR, et al. XRCC3 and XPD/ERCC2 single nucleotide polymorphisms and the risk of cancer: a HuGE review. Am J Epidemiol 2006;164:297–302.
22. Han S, Zhang HT, Wang Z, et al. DNA repair gene XRCC3 polymorphisms and cancer risk: a metaanalysis of 48 case-control studies. Eur J Hum Genet 2006;14:1136–44.
23. Hung RJ, Boffetta P, Canzian F, et al. Sequence variants in cell cycle control pathway, X-ray exposure, and lung cancer risk: a multicenter case-control study in central Europe. Cancer Res 2006;66: 8280–6.
24. Wu X, Zhao H, Amos CI, et al. p53 genotypes and haplotypes associated with lung cancer susceptibility and ethnicity. J Natl Cancer Inst 2002;94:681–90.
25. Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. Nat Genet 2006;38:3–5.
26. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452:633–7.
27. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008;40:616–22.
28. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452:638–42.
29. Cui Y, Morgenstern H, Greenland S, et al. Polymorphism of xeroderma pigmentosum group G and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. Int J Cancer 2006;118:714–20.
30. Zheng YL, Loffredo CA, Yu Z, et al. Bleomycin-induced chromosome breaks as a risk marker for lung cancer: a case-control study with population and hospital controls. Carcinogenesis 2003;24:269–74.
31. Zhou W, Liu G, Miller DP, et al. Polymorphisms in the DNA repair genes XRCC1 and ERCC2, smoking, and lung cancer risk. Cancer Epidemiol Biomarkers Prev 2003;12:359–65.
32. Wichmann H, Gieger C, Illig T; for the MONICA/KORA Study Group. KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 2005;67: S26–30.
33. Popanda O, Schattenberg T, Phong CT, et al. Specific combinations of DNA repair gene variants and increased risk for non-small cell lung cancer. Carcinogenesis 2004;25:2433–41.
34. Kiyohara C, Yamamura KI, Nakanishi Y, Takayama K, Hara N. Polymorphism in GSTM1, GSTT1, and GSTP1 and susceptibility to lung cancer in a Japanese population. Asian Pac J Cancer Prev 2000;1: 293–8.
35. Shen M, Rothman N, Berndt SI, et al. Polymorphisms in folate metabolic genes and lung cancer risk in Xuan Wei, China. Lung Cancer 2005;49:299–309.
36. Seow A, Poh WT, Teh M, et al. Diet, reproductive factors and lung cancer risk among Chinese women in Singapore: evidence for a protective effect of soy in nonsmokers. Int J Cancer 2002; 97:365–71.