Table 1. Epithelial–mesenchymal transition (EMT)-associated markers in clinical samples predict patient prognosis

EMT-associated gene	Characteristics	Cancer types	Reference (author)
Epithelial marker			
E-cadherin	Type I cell–cell adhesion glycoprotein	Breast cancer	Gould Rothberg and Bracken[25]
		Gastric cancer	Chan *et al.*[24]
		Colorectal cancer	Doridi *et al.*[84]
Claudin-1	Tight junctions restrict lateral diffusion of lipids and membrane proteins	Lung cancer	Chao *et al.*[85]
		Renal cell carcinoma	Fritzsche *et al.*[86]
		Ovarian carcinoma	Kleinberg *et al.*[87]
Mesenchymal marker			
Vimentin	Intermediate filaments represent a third class of cytoskeletal elements	Breast cancer	Thomas *et al.*[88]
		Lung cancer	Al-Saad *et al.*[89]
		Gastric cancer	Utsunomiya *et al.*[90]
N-cadherin	Type I cell–cell adhesion glycoprotein	Esophageal cancer	Yoshinaga *et al.*[91]
		Lung cancer	Nakashima *et al.*[92]
		Urothelial tumor	Lascombe *et al.*[93]
Fibronectin	High-molecular weight extracellular matrix glycoprotein	Bladder tumor	Mutlu *et al.*[94]
		Colorectal cancer	Inufusa *et al.*[95]
		Ovarian carcinoma	Franke *et al.*[96]
Transcription factor			
Snail	Zinc finger transcriptional repressor	Adenocortical carcinoma	Waldmann *et al.*[97]
		Esophageal cancer	Natsugoe *et al.*[98]
		Hepatocellular carcinoma	Miyoshi *et al.*[99]
Slug	Zinc finger transcriptional repressor	Lung cancer	Shih *et al.*[100]
		Colorectal cancer	Shioiri *et al.*[101]
		Esophageal cancer	Uchikado *et al.*[102]
Twist	Basic helix-loop-helix transcription factors	Cervical cancer	Shibata *et al.*[103]
		Ovarian carcinoma	Hosono *et al.*[104]
		Breast cancer	Martin *et al.*[105]

metastatic lung adenocarcinoma (with evidence of mutant K-ras and p53) could transit reversibly between epithelial and mesenchymal states, a property that was regulated by the miR-200 family.[44] Furthermore, two recent independent studies showed that members of the miR-200 family can induce the EMT process and regulate the sensitivity to epidermal growth factor receptor (EGFR) in bladder cancer cells and to gemcitabine in pancreatic cancer cells.[45,46] As for regulating TGF-β, microRNAs related to TGF-β signaling such as miR-155 and miR-29a have been identified in breast cancer tissues.[47,48] It is important to identify microRNAs involved in EMT to elucidate up-stream regulators of various known signal pathways.

Microenvironment and EMT

The tumor microenvironment is composed of the extracellular matrix (ECM), cancer-associated fibroblasts, myofibroblasts, immune cells, and soluble factors required for cancer progression and metastasis. Interaction among cancer cells in the tumor microenvironment can induce EMT by auto- and/or paracrine secretion of mediators such as growth factors, cytokines, and ECM proteins.[21] Media conditioned by cultures of cancer-associated fibroblast induce EMT in breast cancer cells.[49] In a comparison of the central areas of primary colorectal cancer and corresponding metastases, nuclear β-catenin was found in dedifferentiated mesenchyme-like tumor cells at the invasive front and it was localized to the membrane and cytoplasm.[50] This study suggested that the tumor microenvironment may induce or maintain EMT (Fig. 2). For instance, cancer-associated fibroblasts may be supplied from cancer cells undergoing EMT.[51] Similarly, oral squamous cancer cells can directly induce a myofibroblastic phenotype via secretion of TGF-β. TGF-β signaling by stromal myofibroblast can induce secretion of hepatocyte growth factor (HGF) which promotes cancer cell proliferation and invasion.[52]

Drug Resistance and EMT

Cells undergoing EMT become invasive and develop resistance to anticancer agents (Fig. 2). In fact, EMT can be induced by anticancer agents, and stress conditions such as exposure to radiation and hypoxic conditions.[53,54] Up-regulation of *TWIST* was associated with cellular resistance to paclitaxel in human nasopharyngeal, bladder, ovarian, and prostate cancers.[55] In colorectal cancer, stable oxaliplatin-resistant cells established by chronic exposure to oxaliplatin can acquire the ability to migrate and invade with phenotypic changes resembling EMT (spindle-cell shape, loss of polarity, intercellular separation, and pseudopodia formation).[56] In pancreatic and ovarian cancer, stable cell lines resistant to gemcitabine and paclitaxel established by continuous exposure can undergo EMT with increased expression of Snail and Twist, EMT-regulatory transcription factors.[57,58]

Various types of molecularly targeted agents have been developed and used against many carcinomas with or without combination of traditional anticancer agents, leading to improved clinical outcome and survival rate.[59,60] However, EMT reportedly confers resistance to these targeted agents. For example, lung cancer cell lines having undergone EMT, expressing vimentin and/or fibronectin, were insensitive to the growth inhibitory effects of EGFR kinase inhibition (erotinib) *in vitro* and in xenografts[61] as well as other EGFR inhibitors such as gefitinib and cetuximab.[62,63] We have often encountered patients who have suffered relapses after drug treatment, even when the tumors were initially highly sensitive. Thus, EMT can lead to resistance to multiple drugs and permit rapid progression of the tumor. These clinical findings may be attributed to the inherent characteristics of EMT. Clarifying the correlation between EMT and drug resistance may help clinicians select an optimal anticancer drug treatment.

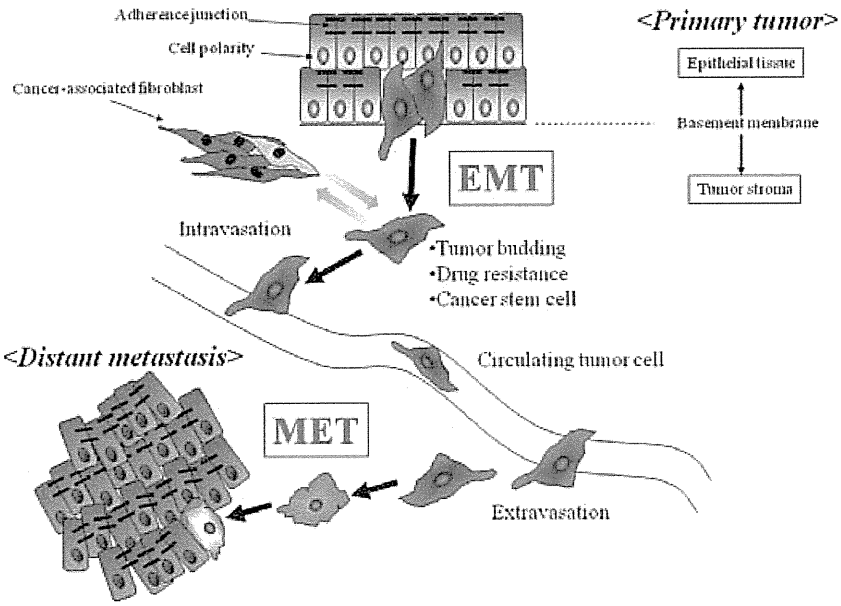

Fig. 2. The epithelial–mesenchymal transition (EMT) and mesenchymal–epithelial transition (MET) are involved in cancer metastasis. Cancer cells undergoing EMT in a primary tumor disseminate through the fragmented basement membrane and acquire the characteristics of drug resistance and cancer stem cells. They can be recognized in tumor buds in histological specimens. EMT cells invade into tumor stroma and enter the circulation, allowing transport to distant organs. At metastatic sites, solitary cancer cells form the new metastatic focus through MET.

Cancer Stem Cells and EMT

Cancer researchers have recently found a minor fraction of cells (cancer stem cells [CSC]) with the ability to self-renew and give rise to differentiated tumor cells. CSC have been identified in breast, colon, and pancreatic cancer.[64–66] CSC as well as cells undergoing EMT are considered to be more resistant to toxic injuries and chemoradiation therapy than differentiated daughter cells.[67,68] Furthermore, cancer cells under hypoxic conditions acquire the properties of CSC.[69,70] Even though evidence indicates a relationship between EMT and cancer cells with the traits of stemness,[71] CSC are rare in whole tumor tissues.[68,72] However, it remains controversial among pathologists whether CSC as well as cells undergoing EMT exist in human cancer tissues.[73] Intriguingly, Mani *et al.* initially disclosed that immortalized human mammary epithelial cells (HMLEs) undergoing EMT are CSC-like as characterized by their $CD44^{high}/CD24^{low}$ phenotype.[16] These investigators induced EMT in HMLEs by ectopic expression of Twist or Snail, known inducers of EMT. The cells undergoing EMT acquired a fibroblastoid mesenchymal appearance. Furthermore, Mani *et al.* observed down-regulation of epithelial markers such as E-cadherin and up-regulation of mesenchymal markers such as N-cadherin, vimentin, and fibronectin. They also noted a $CD44^{high}/CD24^{low}$ expression pattern associated with human breast CSCs. Furthermore, they revealed that the cells undergoing EMT had the properties of CSC, including self-renewal and the capacity to form mammospheres. These findings suggest that EMT may play a role in the development of CSC and properties of invasiveness, metastasis, recurrence, and chemoresistance (Fig. 2).

Clinical Significance of EMT

EMT-associated markers in clinical samples and their effects on prognosis are summarized in Table 1. Most EMT-associated markers have been identified in histological specimens. However, the existence of EMT cells in clinical specimens has been challenged.[74] In response, Voulgari *et al.* suggested that the controversy between experimental and clinical studies is due to the 'spatial' and 'temporal' heterogeneity of EMT (Fig. 3).[19] Cells undergoing EMT may gain metastatic potential but may constitute only a small proportion of the total population of tumor cells. Tumor budding is commonly observed in clinical practice, and it consists of a single cancer cell or small cell cluster at the invasive front of tumor tissues. Indeed, cancer cells in tumor buds have down-regulated E-cadherin[75] and have characteristics of CSC.[76] Therefore, identification of cancer cells undergoing EMT in clinical specimens is difficult for pathologists.

The temporal heterogeneity of EMT (and the reverse, MET) is readily explained. MET is observed *in vitro* following addition of bone morphogenetic protein 7 (BMP7), removal of an EMT-inducer such as TGF-β, and establishment of hypoxic conditions.[54,77] A similar process may occur at metastatic sites which require cancer cells to recover the expression of E-cadherin for cell adhesion. The phenotypes of metastatic specimens are often compared with primary specimens to confirm the diagnosis by hematoxylin–eosin staining. The presence of the same cancer cell characteristics or phenotypes in both primary and metastatic lesions can provide the diagnosis of cancer metastasis. Therefore, the occurrence of MET could make it difficult to prove that EMT, a transient phenomenon that involves only a minority of cells, has occurred in human cancer specimens. However, EMT-associated genes obviously are useful as predictive biomarkers (Table 1). Clinical verification of EMT will require advanced techniques such as *in vivo* imaging.

Treatments Targeting EMT

As shown in Figure 1, EMT-related pathways provide targets for therapy. For instance, inhibition of integrin-linked kinase (ILK) increases the sensitivity of mesenchymal cells to EGFR-target therapy in hepatocellular carcinoma.[63] In *in vitro* studies, Src kinase inhibitors effectively inhibit the growth of cells undergoing EMT.[78] Furthermore, the inhibition of hedgehog signaling can prevent pancreatic cancer cells from acquiring tumor-initiating property and undergoing EMT.[79,80]

RNA interference and microRNA are new technologies in drug development. For instance, silencing of Snail by shRNA induced MET and reduced *in vivo* tumor growth.[81] As for microRNA, Krutzfeldt *et al.* disclosed that specific silencers of endogenous miRNAs, antagomirs, are powerful tools to silence specific miRNAs *in vivo*.[82] Therefore, microRNAs associated with EMT such as the miR-10b and miR-200 family could be exploited as therapeutic strategies in the future.

doi: 10.1111/j.1349-7006.2009.01419.x

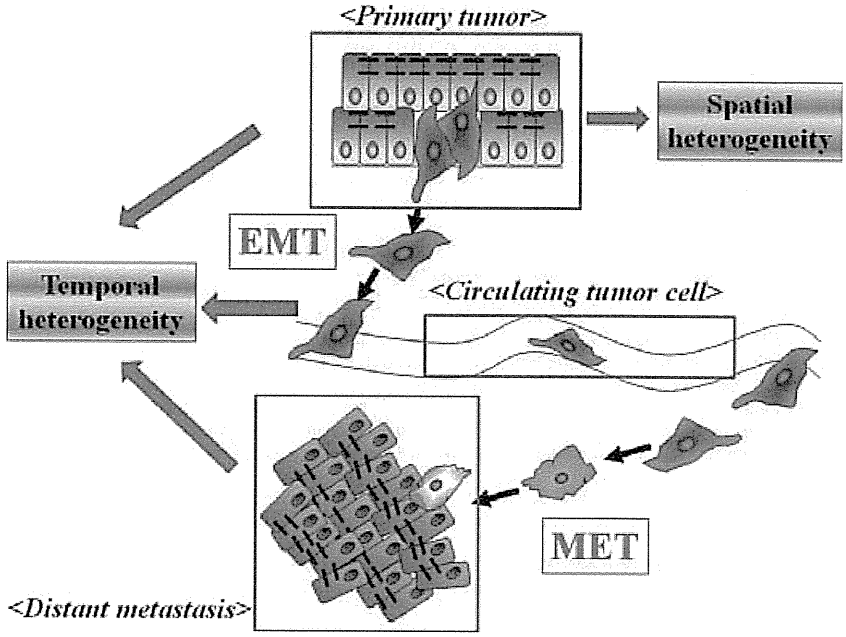

Fig. 3. Spatial and temporal heterogeneity of the epithelial–mesenchymal transition (EMT). Cancer cells undergoing EMT are expected to be only a small proportion of primary tumor tissues. EMT cells transported to metastatic sites are expected to undergo and mesenchymal–epithelial transition (MET). Therefore, the spatial and temporal heterogeneity of EMT/MET severely restricts the ability of pathologists to detect cancer cells undergoing EMT in histological sections.

Furthermore, the tumor microenvironment, which contributes to the maintenance of EMT, could be targeted. A small-interfering RNA targeted at TGF-β reportedly reduces metastasis *in vivo*,[83] and this observation could be applied to TGF-β secreted by tumor stroma. Note that reducing EMT could also lessen the occurrence of anticancer drug resistance and thereby improve the efficacy of conventional therapy. To eradicate cancer cells effectively and cause minimal toxicity to normal cells, further studies are required to define the molecular differences between EMT in embryological development and that in cancer progression.

Perspectives

During the past few decades, an increasing number of studies have shown that EMT is associated with cancer progression, metastasis, and drug resistance. Furthermore, improved understanding of microRNAs and cancer stem cells will clarify the processes underlying EMT. Current understanding of traditional signal pathways coupled with these new concepts could accelerate progress in cancer research. However, the multimodal nature of these complex pathways presents formidable challenges to researchers attempting to inhibit the onset of EMT. Finally, the clinical evidence supporting the role of EMT in cancer progression is still relatively weak. Thus, better methods for EMT detection in patient samples are needed.

Acknowledgments

This work was supported by the following grants and foundations: CREST, Japan Science and Technology Agency (JST); Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research, grant numbers 20390360, 20591547, 20790960, 21591644, 21791295, 21791297, 215921014, and 21679006.

References

1 Weigelt B, Peterse JL, van't Veer LJ. Breast cancer metastasis: markers and models. *Nat Rev Cancer* 2005; **5**: 591–602.
2 Citron ML, Berry DA, Cirrincione C *et al.* Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741. *J Clin Oncol* 2003; **21**: 1431–9.
3 Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2002; **2**: 563–72.
4 Woodhouse EC, Chuaqui RF, Liotta LA. General mechanisms of metastasis. *Cancer* 1997; **80**: 1529–37.
5 Mehes G, Witt A, Kubista E, Ambros PF. Circulating breast cancer cells are frequently apoptotic. *Am J Pathol* 2001; **159**: 17–20.
6 Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* 2009; **9**: 265–73.
7 Thiery JP. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* 2002; **2**: 442–54.
8 Ngan CY, Yamamoto H, Seshimo I *et al.* Quantitative evaluation of vimentin expression in tumour stroma of colorectal cancer. *Br J Cancer* 2007; **96**: 986–92.
9 Raymond WA, Leong AS. Vimentin – a new prognostic parameter in breast carcinoma? *J Pathol* 1989; **158**: 107–14.
10 Dorudi S, Sheffield JP, Poulsom R, Northover JM, Hart IR. E-cadherin expression in colorectal cancer. An immunocytochemical and in situ hybridization study. *Am J Pathol* 1993; **142**: 981–6.
11 Kowalski PJ, Rubin MA, Kleer CG. E-cadherin expression in primary carcinomas of the breast and its distant metastases. *Breast Cancer Res* 2003; **5**: R217–22.
12 Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* 2007; **7**: 415–28.
13 Gregory PA, Bert AG, Paterson EL *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol* 2008; **10**: 593–601.
14 Korpal M, Lee ES, Hu G, Kang Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J Biol Chem* 2008; **283**: 14910–4.
15 Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev* 2008; **22**: 894–907.
16 Mani SA, Guo W, Liao MJ *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008; **133**: 704–15.
17 Greenburg G, Hay ED. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *J Cell Biol* 1982; **95**: 333–9.

18 Hay ED. The mesenchymal cell, its role in the embryo, and the remarkable signaling mechanisms that create it. *Dev Dyn* 2005; **233**: 706–20.
19 Voulgari A, Pintzas A. Epithelial-mesenchymal transition in cancer metastasis: mechanisms, markers and strategies to overcome drug resistance in the clinic. *Biochim Biophys Acta* 2009; **1796**: 75–90.
20 Zavadil J, Bottinger EP. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* 2005; **24**: 5764–74.
21 Moustakas A, Heldin CH. Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. *Cancer Sci* 2007; **98**: 1512–20.
22 Boyer B, Valles AM, Edme N. Induction and regulation of epithelial-mesenchymal transitions. *Biochem Pharmacol* 2000; **60**: 1091–9.
23 Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* 2008; **14**: 818–29.
24 Chan AO, Chu KM, Lam SK *et al.* Soluble E-cadherin is an independent pretherapeutic factor for long-term survival in gastric cancer. *J Clin Oncol* 2003; **21**: 2288–93.
25 Gould Rothberg BE, Bracken MB. E-cadherin immunohistochemical expression as a prognostic factor in infiltrating ductal carcinoma of the breast: a systematic review and meta-analysis. *Breast Cancer Res Treat* 2006; **100**: 139–48.
26 Becker KF, Atkinson MJ, Reich U *et al.* E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer Res* 1994; **54**: 3845–52.
27 Berx G, Cleton-Jansen AM, Nollet F *et al.* E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. *EMBO J* 1995; **14**: 6107–15.
28 Graff JR, Herman JG, Lapidus RG *et al.* E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Res* 1995; **55**: 5195–9.
29 Yoshiura K, Kanai Y, Ochiai A, Shimoyama Y, Sugimura T, Hirohashi S. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc Natl Acad Sci U S A* 1995; **92**: 7416–9.
30 Miettinen PJ, Ebner R, Lopez AR, Derynck R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J Cell Biol* 1994; **127**: 2021–36.
31 Bierie B, Moses HL. Tumour microenvironment: TGFbeta: the molecular Jekyll and Hyde of cancer. *Nat Rev Cancer* 2006; **6**: 506–20.
32 Blobe GC, Schiemann WP, Lodish HF. Role of transforming growth factor beta in human disease. *N Engl J Med* 2000; **342**: 1350–8.
33 Hahn SA, Schutte M, Hoque AT *et al.* DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* 1996; **271**: 350–3.
34 Miyaki M, Iijima T, Konishi M *et al.* Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene* 1999; **18**: 3098–103.
35 Derynck R, Akhurst RJ, Balmain A. TGF-beta signaling in tumor suppression and cancer progression. *Nat Genet* 2001; **29**: 117–29.
36 Batlle E, Sancho E, Franci C *et al.* The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* 2000; **2**: 84–9.
37 Comijn J, Berx G, Vermassen P *et al.* The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell* 2001; **7**: 1267–78.
38 Eger A, Aigner K, Sonderegger S *et al.* DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene* 2005; **24**: 2375–85.
39 Hajra KM, Chen DY, Fearon ER. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res* 2002; **62**: 1613–8.
40 Yang J, Mani SA, Donaher JL *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 2004; **117**: 927–39.
41 Larue L, Bellacosa A. Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene* 2005; **24**: 7443–54.
42 Peinado H, Portillo F, Cano A. Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol* 2004; **48**: 365–75.
43 Ma L, Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* 2007; **449**: 682–8.
44 Gibbons DL, Lin W, Creighton CJ *et al.* Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes Dev* 2009; **23**: 2140–51.
45 Adam L, Zhong M, Choi W *et al.* miR-200 expression regulates epithelial-to-mesenchymal transition in bladder cancer cells and reverses resistance to epidermal growth factor receptor therapy. *Clin Cancer Res* 2009; **15**: 5060–72.
46 Li Y, VandenBoom TG 2nd, Kong D *et al.* Up-regulation of miR-200 and let-7 by natural agents leads to the reversal of epithelial-to-mesenchymal transition in gemcitabine-resistant pancreatic cancer cells. *Cancer Res* 2009; **69**: 6704–12.
47 Gebeshuber CA, Zatloukal K, Martinez J. miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis. *EMBO Rep* 2009; **10**: 400–5.
48 Kong W, Yang H, He L *et al.* MicroRNA-155 is regulated by the transforming growth factor beta/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA. *Mol Cell Biol* 2008; **28**: 6773–84.
49 Lebret SC, Newgreen DF, Thompson EW, Ackland ML. Induction of epithelial to mesenchymal transition in PMC42-LA human breast carcinoma cells by carcinoma-associated fibroblast secreted factors. *Breast Cancer Res* 2007; **9**: R19.
50 Brabletz T, Jung A, Reu S *et al.* Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc Natl Acad Sci U S A* 2001; **98**: 10356–61.
51 Radisky DC, Kenny PA, Bissell MJ. Fibrosis and cancer: do myofibroblasts come also from epithelial cells via EMT? *J Cell Biochem* 2007; **101**: 830–9.
52 Lewis MP, Lygoe KA, Nystrom ML *et al.* Tumour-derived TGF-beta1 modulates myofibroblast differentiation and promotes HGF/SF-dependent invasion of squamous carcinoma cells. *Br J Cancer* 2004; **90**: 822–32.
53 Jung JW, Hwang SY, Hwang JS, Oh ES, Park S, Han IO. Ionising radiation induces changes associated with epithelial-mesenchymal transdifferentiation and increased cell motility of A549 lung epithelial cells. *Eur J Cancer* 2007; **43**: 1214–24.
54 Yang MH, Wu MZ, Chiou SH *et al.* Direct regulation of TWIST by HIF-1alpha promotes metastasis. *Nat Cell Biol* 2008; **10**: 295–305.
55 Wang X, Ling MT, Guan XY *et al.* Identification of a novel function of TWIST, a bHLH protein, in the development of acquired taxol resistance in human cancer cells. *Oncogene* 2004; **23**: 474–82.
56 Yang AD, Fan F, Camp ER *et al.* Chronic oxaliplatin resistance induces epithelial-to-mesenchymal transition in colorectal cancer cell lines. *Clin Cancer Res* 2006; **12**: 4147–53.
57 Kajiyama H, Shibata K, Terauchi M *et al.* Chemoresistance to paclitaxel induces epithelial-mesenchymal transition and enhances metastatic potential for epithelial ovarian carcinoma cells. *Int J Oncol* 2007; **31**: 277–83.
58 Shah AN, Summy JM, Zhang J, Park SI, Parikh NU, Gallick GE. Development and characterization of gemcitabine-resistant pancreatic tumor cells. *Ann Surg Oncol* 2007; **14**: 3629–37.
59 Hurwitz H, Fehrenbacher L, Novotny W *et al.* Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2004; **350**: 2335–42.
60 Piccart-Gebhart MJ, Procter M, Leyland-Jones B *et al.* Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005; **353**: 1659–72.
61 Thomson S, Buck E, Petti F *et al.* Epithelial to mesenchymal transition is a determinant of sensitivity of non-small-cell lung carcinoma cell lines and xenografts to epidermal growth factor receptor inhibition. *Cancer Res* 2005; **65**: 9455–62.
62 Frederick BA, Helfrich BA, Coldren CD *et al.* Epithelial to mesenchymal transition predicts gefitinib resistance in cell lines of head and neck squamous cell carcinoma and non-small cell lung carcinoma. *Mol Cancer Ther* 2007; **6**: 1683–91.
63 Fuchs BC, Fujii T, Dorfman JD *et al.* Epithelial-to-mesenchymal transition and integrin-linked kinase mediate sensitivity to epidermal growth factor receptor inhibition in human hepatoma cells. *Cancer Res* 2008; **68**: 2391–9.
64 Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 2003; **100**: 3983–8.
65 Li C, Heidt DG, Dalerba P *et al.* Identification of pancreatic cancer stem cells. *Cancer Res* 2007; **67**: 1030–7.
66 O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 2007; **445**: 106–10.
67 Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature* 2001; **414**: 105–11.
68 Sagar J, Chaib B, Sales K, Winslet M, Seifalian A. Role of stem cells in cancer therapy and cancer stem cells: a review. *Cancer Cell Int* 2007; **7**: 9.
69 Axelson H, Fredlund E, Ovenberger M, Landberg G, Pahlman S. Hypoxia-induced dedifferentiation of tumor cells – a mechanism behind heterogeneity and aggressiveness of solid tumors. *Semin Cell Dev Biol* 2005; **16**: 554–63.
70 Platet N, Liu SY, Atifi ME *et al.* Influence of oxygen tension on CD133 phenotype in human glioma cell cultures. *Cancer Lett* 2007; **258**: 286–90.
71 Prindull G. Hypothesis: cell plasticity, linking embryonal stem cells to adult stem cell reservoirs and metastatic cancer cells? *Exp Hematol* 2005; **33**: 738–46.
72 Ishii H, Iwatsuki M, Ieta K *et al.* Cancer stem cells and chemoradiation resistance. *Cancer Sci* 2008; **99**: 1871–7.
73 Christiansen JJ, Rajasekaran AK. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer Res* 2006; **66**: 8319–26.
74 Garber K. Epithelial-to-mesenchymal transition is important to metastasis, but questions remain. *J Natl Cancer Inst* 2008; **100**: 232–3. 9.

doi: 10.1111/j.1349-7006.2009.01419.x

75 Masaki T, Goto A, Sugiyama M *et al.* Possible contribution of CD44 variant 6 and nuclear beta-catenin expression to the formation of budding tumor cells in patients with T1 colorectal carcinoma. *Cancer* 2001; **92**: 2539–46.
76 Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells – an integrated concept of malignant tumour progression. *Nat Rev Cancer* 2005; **5**: 744–9.
77 Rees JR, Onwuegbusi BA, Save VE, Alderson D, Fitzgerald RC. In vivo and in vitro evidence for transforming growth factor-beta1-mediated epithelial to mesenchymal transition in esophageal adenocarcinoma. *Cancer Res* 2006; **66**: 9583–90.
78 Finn RS, Dering J, Ginther C *et al.* Dasatinib, an orally active small molecule inhibitor of both the src and abl kinases, selectively inhibits growth of basal-type/''triple-negative'' breast cancer cell lines growing in vitro. *Breast Cancer Res Treat* 2007; **105**: 319–26.
79 Feldmann G, Fendrich V, McGovern K *et al.* An orally bioavailable small-molecule inhibitor of Hedgehog signaling inhibits tumor initiation and metastasis in pancreatic cancer. *Mol Cancer Ther* 2008; **7**: 2725–35.
80 Feldmann G, Dhara S, Fendrich V *et al.* Blockade of hedgehog signaling inhibits pancreatic cancer invasion and metastases: a new paradigm for combination therapy in solid cancers. *Cancer Res* 2007; **67**: 2187–96.
81 Olmeda D, Jorda M, Peinado H, Fabra A, Cano A. Snail silencing effectively suppresses tumour growth and invasiveness. *Oncogene* 2007; **26**: 1862–74.
82 Krutzfeldt J, Rajewsky N, Braich R *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 2005; **438**: 685–9.
83 Moore LD, Isayeva T, Siegal GP, Ponnazhagan S. Silencing of transforming growth factor-beta1 in situ by RNA interference for breast cancer: implications for proliferation and migration in vitro and metastasis in vivo. *Clin Cancer Res* 2008; **14**: 4961–70.
84 Dorudi S, Hanby AM, Poulsom R, Northover J, Hart IR. Level of expression of E-cadherin mRNA in colorectal cancer correlates with clinical outcome. *Br J Cancer* 1995; **71**: 614–6.
85 Chao YC, Pan SH, Yang SC, *et al.* Claudin-1 is a metastasis suppressor and correlates with clinical outcome in lung adenocarcinoma. *Am J Respir Crit Care Med* 2009; **179**: 123–33.
86 Fritzsche FR, Oelrich B, Johannsen M, *et al.* Claudin-1 protein expression is a prognostic marker of patient survival in renal cell carcinomas. *Clin Cancer Res* 2008; **14**: 7035–42.
87 Kleinberg L, Holth A, Trope CG, Reich R, Davidson B. Claudin upregulation in ovarian carcinoma effusions is associated with poor survival. *Hum Pathol* 2008; **39**: 747–57.
88 Thomas PA, Kirschmann DA, Cerhan JR, *et al.* Association between keratin and vimentin expression, malignant phenotype, and survival in postmenopausal breast cancer patients. *Clin Cancer Res* 1999; **5**: 2698–703.
89 Al-Saad S, Al-Shibli K, Donnem T, Persson M, Bremnes RM, Busund LT. The prognostic impact of NF-kappaB p105, vimentin, E-cadherin and Par6 expression in epithelial and stromal compartment in non-small-cell lung cancer. *Br J Cancer* 2008; **99**: 1476–83.
90 Utsunomiya T, Yao T, Masuda K, Tsuneyoshi M. Vimentin-positive adenocarcinomas of the stomach: co-expression of vimentin and cytokeratin. *Histopathology* 1996; **29**: 507–16.
91 Yoshinaga K, Inoue H, Utsunomiya T, *et al.* N-cadherin is regulated by activin A and associated with tumor aggressiveness in esophageal carcinoma. *Clin Cancer Res* 2004; **10**: 5702–7.
92 Nakashima T, Huang C, Liu D, *et al.* Neural-cadherin expression associated with angiogenesis in non-small-cell lung cancer patients. *Br J Cancer* 2003; **88**: 1727–33.
93 Lascombe I, Clairotte A, Fauconnet S, *et al.* N-cadherin as a novel prognostic marker of progression in superficial urothelial tumors. *Clin Cancer Res* 2006; **12**: 2780–7.
94 Mutlu N, Turkeri L, Emerk K. Analytical and clinical evaluation of a new urinary tumor marker: bladder tumor fibronectin in diagnosis and follow-up of bladder cancer. *Clin Chem Lab Med* 2003; **41**: 1069–74.
95 Inufusa H, Nakamura M, Adachi T, *et al.* Localization of oncofetal and normal fibronectin in colorectal cancer. Correlation with histologic grade, liver metastasis, and prognosis. *Cancer* 1995; **75**: 2802–8.
96 Franke FE, Von Georgi R, Zygmunt M, Munstedt K. Association between fibronectin expression and prognosis in ovarian carcinoma. *Anticancer Res* 2003; **23**: 4261–7.
97 Waldmann J, Feldmann G, Slater EP, *et al.* Expression of the zinc-finger transcription factor Snail in adrenocortical carcinoma is associated with decreased survival. *Br J Cancer* 2008; **99**: 1900–7.
98 Natsugoe S, Uchikado Y, Okumura H, *et al.* Snail plays a key role in E-cadherin-preserved esophageal squamous cell carcinoma. *Oncol Rep* 2007; **17**: 517–23.
99 Miyoshi A, Kitajima Y, Kido S, *et al.* Snail accelerates cancer invasion by upregulating MMP expression and is associated with poor prognosis of hepatocellular carcinoma. *Br J Cancer* 2005; **92**: 252–8.
100 Shih JY, Tsai MF, Chang TH, *et al.* Transcription repressor slug promotes carcinoma invasion and predicts outcome of patients with lung adenocarcinoma. *Clin Cancer Res* 2005; **11**: 8070–8.
101 Shioiri M, Shida T, Koda K, *et al.* Slug expression is an independent prognostic parameter for poor survival in colorectal carcinoma patients. *Br J Cancer* 2006; **94**: 1816–22.
102 Uchikado Y, Natsugoe S, Okumura H, *et al.* Slug Expression in the E-cadherin preserved tumors is related to prognosis in patients with esophageal squamous cell carcinoma. *Clin Cancer Res* 2005; **11**: 1174–80.
103 Shibata K, Kajiyama H, Ino K, *et al.* Twist expression in patients with cervical cancer is associated with poor disease outcome. *Ann Oncol* 2008; **19**: 81–5.
104 Hosono S, Kajiyama H, Terauchi M, *et al.* Expression of Twist increases the risk for recurrence and for poor survival in epithelial ovarian carcinoma patients. *Br J Cancer* 2007; **96**: 314–20.
105 Martin TA, Goyal A, Watkins G, Jiang WG. Expression of the transcription factors snail, slug, and twist and their clinical significance in human breast cancer. *Ann Surg Oncol* 2005; **12**: 488–96.

Carcinogenesis vol.31 no.4 pp.712–718, 2010
doi:10.1093/carcin/bgq010
Advance Access publication January 18, 2010

Plasma levels of C-reactive protein and serum amyloid A and gastric cancer in a nested case–control study: Japan Public Health Center-based prospective study

Shizuka Sasazuki*, Manami Inoue, Norie Sawada, Motoki Iwasaki, Taichi Shimazu, Taiki Yamaji, Shoichiro Tsugane for the Japan Public Health Center-Based Prospective Study Group

Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo 104-0045, Japan.

*To whom correspondence should be addressed.
Tel: +81 3 3542 2511 ext. 3378; Fax: +81 3 3547 8578;
Email: ssasazuk@gan2.res.ncc.go.jp

Gastric carcinogenesis may be under the combined influence of factors related to the host, *Helicobacter pylori* bacterial virulence and the environment. One possible host-related factor is the inflammatory or immune response. To clarify this point, we investigated the association between plasma levels of C-reactive protein (CRP) and serum amyloid A (SAA) and the subsequent risk of gastric cancer in a population-based nested case–control study. Subjects were observed from 1990 to 2004. Among 36 745 subjects who answered the baseline questionnaire and provided blood samples, 494 gastric cancer cases were identified and matched to 494 controls for our analysis. The overall distribution of CRP and SAA was not apparently associated with the development of gastric cancer. However, a statistically significant increased risk was observed when subjects were categorized dichotomously. The adjusted odds ratio (OR) for the development of gastric cancer for the CRP-positive group (CRP > 0.18 mg/dl) compared with the CRP-negative group was 1.90 [95% confidence interval (CI): 1.19–3.02, $P = 0.007$]. The OR for the SAA-positive group (SAA > 8 μg/ml) compared with the SAA-negative group was 1.93 (95% CI: 1.22–3.07, $P = 0.005$). In conclusion, our results suggest that those who react strongly to inflammation or who have a high host immune response, as reflected by extremely elevated plasma levels of CRP and SAA, are at a high risk to develop gastric cancer.

Introduction

It is well established that cancer arises in chronically inflamed tissue, and one of the classic examples is *Helicobacter pylori*-associated gastric cancer (1). *Helicobacter pylori* persistently colonizes the gastric mucosa, leading to chronic inflammation, atrophic gastritis and, finally, gastric cancer. There are high interindividual differences in the extent of gastric inflammation among *H.pylori*-infected subjects, and only a small proportion of them develop clinical consequences. This indicates that gastric carcinogenesis may be under the combined influence of factors related to the host, bacterial virulence and the environment. One possible host-related factor is the inflammatory or immune response. Many studies have reported an association between serum proinflammatory cytokines [e.g. interleukin (IL)-6, IL-8 and IL-1β] levels (2–4) or polymorphisms (such as IL-1, IL-2 and IL-8) and gastric cancer risk (5–8), but the results are controversial. The lack of consensus may be partly due to the nature of cytokines, which are components of a large, complex signaling network, and difficulties in measuring their levels and interactions. Measurement of cytokines in plasma is difficult because of their short plasma half-lives and the presence of blocking factors (9). Additionally, combinations of cytokines have been found to have additive, inhibitory or synergistic effects. Therefore, more useful or systematic indicators of host inflammatory or immune response are needed.

C-reactive protein (CRP) is a well-established indicator of inflammation in the body (10). It is an acute-phase reactant that reflects low-grade systemic inflammation and has been studied in a variety of cardiovascular diseases. CRP production by the liver is regulated by cytokines, principally IL-6 and tumor necrosis factor α, which is the main trigger for the production of IL-6 by a variety of cells. In fact, strong positive associations between IL-6, tumor necrosis factor α and CRP were observed (11). Serum amyloid A (SAA) is another major acute-phase reactant. It is a putative serum precursor of the amyloid A protein, which constitutes amyloid fibrils in secondary amyloidosis and is an apolipoprotein associated with the high density lipoprotein 3 fraction of serum (12). In most studies, a parallel increase of SAA and CRP has been observed, although some studies have delineated acute-phase SAA as the more sensitive parameter (13,14). Therefore, to indicate the host inflammatory or immune response systematically, CRP and SAA may be useful markers.

In this large-scale nested case–control study, we aimed to examine whether the host inflammatory or immune response has any association with the development of gastric cancer. To clarify this point, we explored the relation of plasma levels of CRP and SAA to risk of developing gastric cancer. As far as we know, this is the first study to prospectively seek this association in a population.

Abbreviations: BMI, body mass index; CagA, cytotoxin-associated gene A; CI, confidence interval; CRP, C-reactive protein; ICD-O, International Classification of Diseases for Oncology; Ig, immunoglobulin; IL, interleukin; JPHC, Japan Public Health Center; OR, odds ratio; PG, pepsinogen; PHC, public health center; SAA, serum amyloid A.

Materials and methods

Study population

The Japan Public Health Center-based prospective study (JPHC Study) is an ongoing cohort study to investigate cancer, cardiovascular disease and other lifestyle-related diseases. The first group (Cohort I) of the JPHC Study was started in 1990 and the second group (Cohort II) in 1993 (15). The JPHC Study included 140 420 subjects (68 722 men and 71 698 women), defined as all inhabitants in the study areas [27 cities, towns or villages served by 11 public health centers (PHCs)] who were 40–59 years old (Cohort I) or 40–69 years old (Cohort II). Among the study subjects, those registered at one PHC area in Cohort I were excluded from the present analysis because data on cancer incidence were not available. Additionally, one subcohort in Cohort II was excluded because the selection of subjects differed from that of other cohort subjects, i.e. random sampling of residents from a municipality population registry for one city, stratified by 10 year age–gender groups. We thus defined 123 576 subjects (61 009 men and 62 567 women) for the present study. The JPHC Study was approved by the institutional review board of the National Cancer Center, Tokyo, Japan.

Baseline survey

In 1990 for Cohort I and in 1993–1994 for Cohort II, subjects were asked to reply to a lifestyle questionnaire that covered sociodemographic characteristics, medical history, smoking and drinking habits, diet and so on. Details of the food frequency questionnaire included in the baseline survey have been described previously (16). A total of 99 808 (81%) subjects—47 525 men and 52 283 women—responded to the questionnaires.

We excluded subjects who self-reported cancer at baseline ($n = 2136$), those who were not Japanese ($n = 18$) and those who were later discovered to have moved away at baseline ($n = 11$). This left 97 644 eligible subjects (46 803 men and 50 841 women). Among them, 36 745 subjects (38%; 13 467 men and 23 278 women) donated blood samples at health checkups conducted by the PHC in each area. Each subject voluntarily provided 10 ml of blood during the health checkups. As customary, subjects were asked to avoid having a meal later than 21:00 on the day before the examination. The last time of either consuming a meal or drinking water or tea was recorded. The plasma and buffy layer were divided into four tubes, with each tube holding 1.0 ml (3 tubes for plasma and 1 for the buffy layer) and stored at 80°C. Blood was collected from 1990 to 1992 in Cohort I and from 1993 to 1995 in Cohort II.

Follow-up and identification of gastric cancer

In Japan, at the time the study was conducted, a PHC played a role as an organization that provided primary health care, including health checkups,

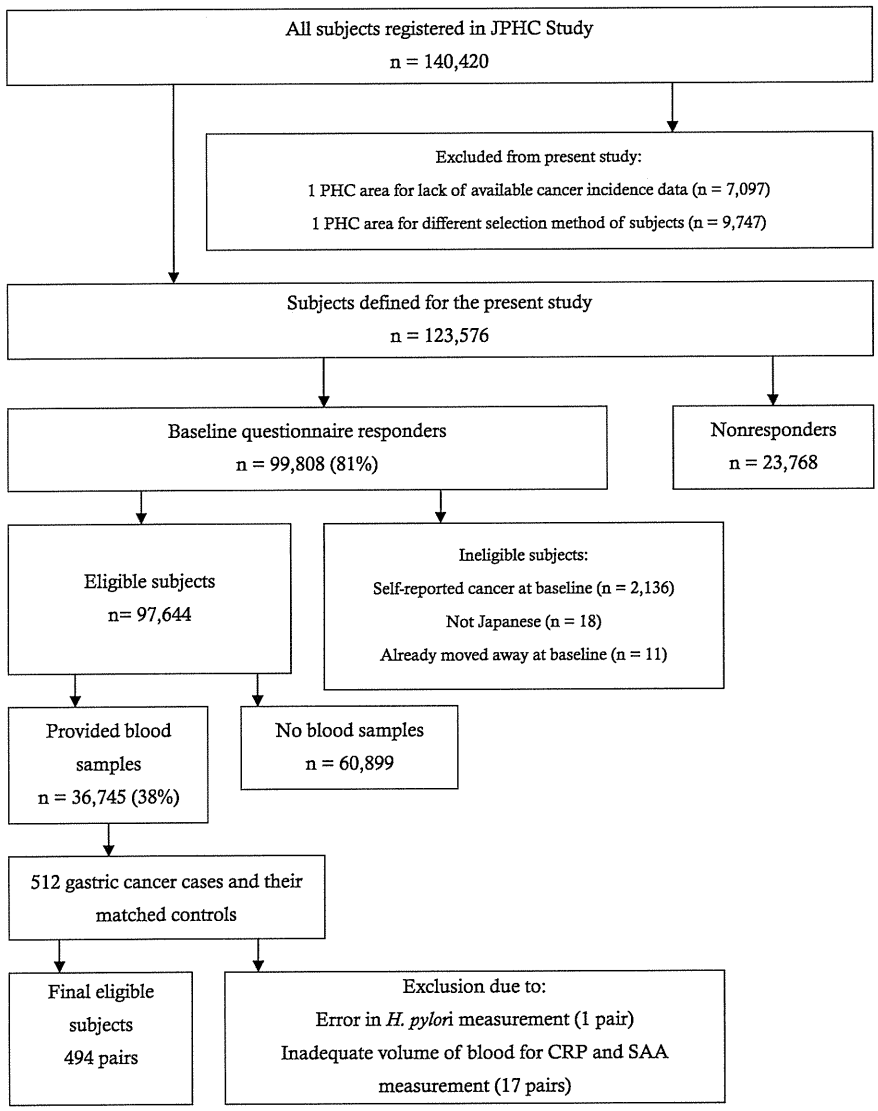

Fig. 1. Flow of study population.

or other health promotion activities for all inhabitants of the municipalities supervised by the PHC. In this study, the main role of the PHC was to collect and report data on mortality, relocation and cancer cases.

Death and relocation

We observed study subjects until 31 December 2004. The changes in residency status, including death, were identified annually through the residential registry in each area. To confirm causes of death, we used mortality data from the Ministry of Health, Labour and Welfare. Residence and death registration are required by law in Japan, and the registries are believed to be complete. Among 36 745 study subjects, 1423 (3.9%) moved away from the study area, 1610 (4.4%) died and 11 (0.03%) were lost to follow-up within the study period.

Cancer registry for JPHC Study

Data on newly diagnosed cases of cancer were collected from two sources: active patient notification from the local major hospitals in the study area and data linkage with population-based registries (usually prefecture-wide). Death certificate information was used as a supplementary information source. In our cancer registry system, the proportion of cases of gastric cancer for which information was based on death certificate notification was 7.6% and on information available from death certificates only was 2.1%. This level of quality for the information was considered satisfactory for the present study.

Identification of gastric cancer and selection of control subjects

Cases of gastric cancer were extracted from the cancer registry for the JPHC Study on the basis of site [International Classification of Diseases for Oncology (ICD-O) code C160–169] (17). Up to the end of the study period, 512 new gastric cancer cases were identified. Until quite recently in Japan, the upper third of the stomach has been called the 'cardia' on the basis of the guidelines for gastric cancer classification (18). Because it seemed difficult to distinguish the cardia, which is mainly located in the esophagogastric junction, from the upper third of the stomach, we combined tumors at these sites into one group for analysis (ICD-O code C160–161). A tumor located on the lower side of the stomach was classified as distal gastric cancer (ICD-O code C162–167). Subsites that could not be classified because of a diffuse lesion (ICD-O code C168) or those with no information (ICD-O code C169) were categorized as an unclassified subsite. Histologic classification was based on one author's (S.S.) review, in consultation with a pathologist, of the record reported by each hospital. The subdivisions were made on the basis of a classification derived by Lauren (19). For each case, one control was selected from subjects who had no history of gastric cancer and who lived in the study area when the case was

diagnosed. Each control was matched to a case for gender, age (±3 years), PHC area, blood donation date (±2 months) and fasting time at blood donation (±5 h). Because of a technical error in measurement of *H.pylori* and inadequate volume of blood available for CRP and SAA measurements, 1 case with its matched control and another 17 pairs (8 cases with their matched controls and 10 controls with their matched cases) were excluded. Finally, we had 494 sets each of cases and controls for use in the present analysis. A flowchart of the study subjects is provided in Figure 1.

Laboratory analysis

CRP and SAA concentrations were determined by the latex agglutination nephelometric immunoassay test (LZ test 'Eiken' CRP-HG; Eiken Kagaku Co. Ltd, Tokyo, Japan; and LZ test 'Eiken' SAA; Eiken Kagaku Co. Ltd, respectively). For the CRP test, based on 10 replicated measurements of three concentrations of blood samples (0.07, 0.50 and 4.41 mg/dl) at the time of analyses, the coefficients of variation were 1.69%, 0.59% and 0.76%, respectively. For SAA, 10 replicated measurements of two concentrations of blood samples (22 μg/ml and 110 μg/ml) yielded a coefficient of variation of reproducibility values of 1.53% and 1.17%. Normal values for the examined parameters were <0.18 mg/dl for CRP and <8 μg/ml for SAA according to the kit's protocol. Both cutoff values were based on data from reports for the same kit. The cutoff value of CRP was set by the iterative truncation method among 478 health checkup samples (20). In brief, after repeated deletion of outliers, mean ± 1.96 SD was considered the normal range. For SAA, after being converted to a logarithm, the value was set as the upper 95th percentile of the distribution of 1056 normal subjects (0–70 years old) (21).

Immunoglobulin (Ig) G antibodies to *H.pylori* were measured with a direct enzyme-linked immunosorbent assay kit (E Plate 'Eiken' *H.pylori* Antibody; Eiken Kagaku Co. Ltd). Levels of IgG were categorized as seropositive and seronegative for *H.pylori* according to a selective cutoff value (≤10 or >10). The cutoff value was based on the results of sensitivity and specificity calculated with the urea test, which is the gold standard (report by company). Assays of cytotoxin-associated gene A (CagA) were performed with the use of an enzyme-linked immunosorbent assay kit, in which horseradish peroxidase was used as the enzyme tracer (CagA IgG EIA; Sceti Co. Ltd, Rome, Italy). According to the manufacturer's protocol, samples with IgG values ≤10 RU/ml must be considered non-reactive for anti-CagA IgG antibodies; samples with IgG values within 10–15 RU/ml must be considered weakly reactive and samples with IgG values >15 RU/ml must be considered reactive for anti-CagA IgG antibodies. With regard to interpretation of these results, reactive and/or questionable samples are considered positive for anti-CagA IgG antibodies, i.e. values >10 are regarded as CagA positive. Serum levels of pepsinogen I and II (PGI and PGII, respectively) were measured by commercial kits based on a two-step enzyme immunoassay (E Plate 'Eiken' Pepsinogen I; Eiken Kagaku Co. Ltd; and E Plate 'Eiken' Pepsinogen II; Eiken Kagaku Co. Ltd). Results were defined as 'atrophic' when the criteria of both PGI level ≤70 ng/ml and PGI : PGII ratio ≤3.0 were fulfilled. Comparing the PG levels between gastric cancer cases and healthy controls retrospectively, Miki (22) reported that applying a PGI level ≤70 ng/ml and a PGI : PGII ratio ≤3.0 as cutoff values was most effective in distinguishing cases from controls. Using these criteria, other authors have showed an extremely high correlation (r = 0.999) between atrophy and age-adjusted gastric cancer mortality among inhabitants of five areas in Japan (23). Among atrophic cases, more severe cases with a PGI level ≤30 ng/ml and PGI : PGII ratio ≤2.0 were defined as severe atrophy.

All measurements were conducted by a person blinded to the case–control situation.

Statistical analysis

Statistical analysis included chi-square test, analysis of variance, analysis of covariance and conditional logistic model. Multiple conditional logistic regression analyses were conducted to control for potential confounding factors. For cardia cancer, smoking status, alcohol consumption (for SAA analysis), intake of salt, body mass index (BMI), family history of gastric cancer, history of infectious or inflammatory disease (i.e. cardiovascular disease, ischemic heart disease, liver disease and kidney disease) and current use of analgesics for lumbago, neuralgia, common cold, arthrosis and joint pain were controlled. For all gastric cancer, all non-cardia cancer, differentiated-type non-cardia cancer and undifferentiated-type non-cardia cancer further adjustment was applied for *H.pylori* infection, atrophy and CagA seropositivity. Smoking status was divided into four groups: never smoker, past smoker, current smoker with <20 cigarettes per day and current smoker with ≥20 cigarettes per day). Alcohol consumption was defined as drinker (>1 day/week) and non-drinker (<1 day/week). BMI was categorized into three groups so that each category included an approximately equal number of controls. Salt was treated as a continuous variable. Family history of gastric cancer was regarded as positive if at least one parent or sibling had gastric cancer. CRP and SAA status (positive/negative) were determined according to the protocol's normal value. Additionally, the non-linear continuous models of the association between CRP and SAA and gastric cancer risk were tested by PROC GAM. Odds ratios (ORs) were calculated relative to the cutoff points of CRP and SAA. Because the distribution was skewed, log transformation was conducted for CRP, SAA, *H.pylori* titer, CagA titer, PGI level and PGII and PGI : PGII ratio, which altered the distribution close to normal in comparisons of the mean values between groups.

Reported P-values were two sided, and all statistical analyses were done with SAS software version 9.1 (SAS Institute Inc., Cary, NC).

Results

Baseline characteristics of cases and controls are shown in Table I. Among listed factors, predominance of *H.pylori* positivity, CagA status, atrophy and family history of gastric cancer were apparent in cases compared with controls.

Table II summarizes the distribution of lifestyle factors and plasma biomarkers according to the CRP and SAA status among controls. Forty-seven (9.5%) and 63 (12.8%) subjects met the criteria for being positive for plasma CRP and SAA, respectively. For CRP status, no factors were differently distributed other than SAA levels; the mean value of SAA among CRP-positive subjects was >10 times that of CRP-negative subjects ($P < 0.0001$). Plasma CRP level among SAA-positive subjects was 13 times that among SAA-negative subjects ($P < 0.0001$). Correlation of the log-transformed CRP and SAA was 0.55 ($P < 0.0001$). Mean daily salt intake was higher in SAA-negative subjects compared with SAA-positive subjects. This may be due to the predominance of male gender and alcohol consumption among SAA-negative subjects, which contribute to high salt intake. When gender and alcohol consumption were adjusted (analysis of covariance), the difference in salt intake was no longer significant ($P = 0.40$). Compared with positive subjects, SAA-negative subjects had a significantly higher *H.pylori* titer against IgG antibody and more frequent distribution of male gender, alcohol consumption, *H.pylori* positivity and atrophy.

Table I. Baseline characteristics of cases and controls

	Case	Control	P-value[a]
n	494	494	
Age	57.3 (0.3)	57.3 (0.3)	Matching value
Men (%)	329 (66.6%)	329 (66.6%)	Matching value
Cigarette smoking			
Never smoker (%)	228 (46.2%)	245 (49.6%)	
Past smoker (%)	91 (18.4%)	98 (19.8%)	
Current smoker with <20 cigarettes per day (%)	133 (26.9%)	109 (22.1%)	
Current smoker with ≥20 cigarettes per day (%)	42 (8.5%)	42 (8.5%)	0.35
Alcohol consumption			
Never or occasional (%)	245 (49.6%)	244 (49.4%)	
≥1 day, <300 g/week (%)	187 (37.9%)	203 (41.1%)	
≥1 day, ≥300 g/week (%)	62 (12.6%)	47 (9.5%)	0.26
BMI			
<25	396 (80.2%)	369 (74.7%)	
25–29.9	89 (18.0%)	113 (22.9%)	
≥30	9 (1.8%)	12 (2.4%)	0.12
Family history of gastric cancer (%)	60 (12.2%)	40 (8.1%)	0.03
Salt (g/day)	5.3 (0.1)	5.1 (0.1)	0.40
Helicobacter pylori positive (%)[b]	463 (93.7%)	371 (75.1%)	<0.0001
Helicobacter pylori positive (%)[c]	489 (99.0%)	445 (90.1%)	<0.0001
CagA (+) (%)	375 (75.9%)	346 (70.0%)	0.04
Atrophy (%)	406 (82.2%)	285 (57.7%)	<0.0001

Values are mean (SE) except where specified otherwise.
[a]Based on chi-square test or analysis of variance.
[b]Based on IgG antibody.
[c]Based on CagA positive and/or *Helicobacter pylori* IgG antibody positive.

Table II. Distribution of lifestyle factors and plasma biomarkers according to CRP and SAA status among control

	CRP status			SAA status		
	Negative (CRP ≤ 0.18 mg/dl)	Positive (CRP > 0.18 mg/dl)	*P*-value[a]	Negative (SAA ≤ 8 μg/ml)	Positive (SAA > 8 μg/ml)	*P*-value[a]
n	447	47		431	63	
Age	57.1 (0.3)	58.6 (1.1)	0.20	57.2 (0.3)	58.1 (0.9)	0.35
Men (%)	296 (66.2%)	33 (70.2%)	0.58	296 (68.7%)	33 (52.4%)	0.01
BMI						
<25	336 (75.2%)	33 (70.2%)		327 (75.9%)	42 (66.7%)	
25–29.9	100 (22.4%)	13 (27.7%)		93 (21.6%)	20 (31.8%)	
≥30	11 (2.5%)	1 (2.1%)	0.71	11 (2.6%)	1 (1.6%)	0.19
Cigarette smoking						
Never smoker (%)	225 (50.3%)	20 (42.6%)		207 (48.0%)	38 (60.3%)	
Past smoker (%)	91 (20.4%)	7 (14.9%)		90 (20.9%)	8 (12.7%)	
Current smoker with <20 cigarettes per day (%)	97 (21.7%)	12 (25.5%)		98 (22.7%)	11 (17.5%)	
Current smoker with ≥20 cigarettes per day (%)	34 (7.6%)	8 (17.0%)	0.12	36 (8.4%)	6 (9.5%)	0.23
Alcohol consumption						
Never or occasional (%)	218 (48.8%)	26 (55.3%)		204 (47.3%)	40 (63.5%)	
≥1 day, <300 g/week (%)	184 (41.2%)	19 (40.4%)		185 (42.9%)	18 (28.6%)	
≥1 day, ≥300 g/week (%)	45 (10.1%)	2 (4.3%)	0.39	42 (9.7%)	5 (7.9%)	0.05
Family history of gastric cancer (%)	36 (8.1%)	4 (8.5%)	0.91	36 (8.4%)	4 (6.4%)	0.85
Salt (g/day)	5.2 (0.1)	4.9 (0.3)	0.43	5.2 (0.1)	4.6 (0.3)	0.04
CRP (mg/dl)/SAA (μg/ml)[b]	3.6 (1.8)	38.6 (5.6)	<0.0001[a]	0.05 (0.03)	0.65 (0.07)	<0.0001[a]
Helicobacter pylori positive (%)[c]	338 (75.6%)	33 (70.2%)	0.42	332 (77.0%)	39 (61.9%)	0.01
Helicobacter pylori positive (%)[d]	403 (90.2%)	42 (89.4%)	0.86	390 (90.5%)	55 (87.3%)	0.43
Helicobacter pylori titer	43.9 (2.3)	36.1 (7.1)	0.31[e]	44.0 (2.3)	37.1 (6.1)	0.02[e]
CagA (+) (%)	314 (70.3%)	32 (68.1%)	0.76	302 (70.1%)	44 (69.8%)	0.97
CagA titer	85.1 (4.2)	74.7 (12.9)	0.72[e]	84.6 (4.3)	80.8 (11.1)	0.82[e]
PGI	28.6 (0.8)	29.7 (2.5)	0.55[e]	28.5 (0.8)	30.1 (2.1)	0.52[e]
PGII	11.2 (0.3)	10.8 (1.0)	0.60[e]	11.2 (0.3)	11.0 (0.8)	0.71[e]
PGI : PGII	3.5 (0.6)	2.9 (1.8)	0.83[e]	3.5 (0.6)	3.2 (1.6)	0.28[e]
Atrophy (%)	260 (58.2%)	25 (53.2%)	0.51	256 (59.4%)	29 (46.0%)	0.04
Severe atrophy (%)	122 (27.3%)	9 (19.2%)	0.23	119 (27.6%)	12 (19.1%)	0.15

Values are mean (SE) except where specified otherwise.
[a]Based on chi-square test or analysis of variance.
[b]Mean plasma CRP level for SAA status and mean plasma SAA level for CRP status.
[c]Based on IgG antibody.
[d]Based on CagA positive and/or *Helicobacter pylori* IgG antibody positive.
[e]Based on analysis of variance of log biomarkers.

In Table III, ORs and 95% confidence intervals (CIs) of CRP positivity for development of gastric cancer are presented by tumor subsite and histologic types. CRP ranged from 0 to 19.1 mg/dl (mean: 0.14 mg/dl, median: 0.033 mg/dl) among cases and from 0 to 9.3 mg/dl (mean: 0.13 mg/dl, median: 0.032 mg/dl) among controls. The risk of developing gastric cancer increased by ~36% among those who were CRP positive; the crude OR equaled 1.36 (95% CI: 0.91–2.02, $P = 0.13$), although with no significance. After being adjusted for potential confounding variables, the point estimate altered substantially and reached the level of statistical significance; the adjusted OR equals 1.90 (95% CI: 1.19–3.02, $P = 0.007$). Among the adjusted covariates, *H.pylori* infection contributed the most to the elevation of risk; adding only *H.pylori* infection to the model elevated the OR to 1.67, which was much higher than the OR for adding CagA seropositivity (adjusted OR = 1.39), atrophy (adjusted OR = 1.48) or even all other lifestyle factors [i.e. cigarette smoking, BMI, family history, history of infectious or inflammatory disease, current drug use of analgesics and salt intake (adjusted OR = 1.47)]. When the cancers were stratified by tumor location and histologic type, the largest OR was demonstrated for cardia cancers, but it failed to reach statistical significance; adjusted OR equaled 3.14 (95% CI: 0.51–19.39, $P = 0.22$). Among non-cardia cancers, the association did not differ much by histologic type. When the analyses were repeated with subjects divided into quartiles according to control distribution of the CRP level (<0.012, 0.012–0.032, 0.032–0.081 and ≥0.081 mg/dl), no apparent association was observed. Compared with the lowest (reference) group, the adjusted ORs (95% CIs) for development of gastric cancer for the second, the third and the highest group were 0.85 (0.56–1.29), 0.96 (0.62–1.47) and 1.35 (0.88–2.07), respectively (P for trend = 0.0496). When CRP was treated as a continuous measure, the adjusted OR for development of gastric cancer was 1.06 (0.87–1.28), for 1 mg/dl increase of log-transformed CRP. Furthermore, non-linear continuous models did not reveal any evidence of dose response.

SAA among cases and controls ranged from 0 to 319.7 μg/ml (mean: 5.9 μg/ml, median: 2.6 μg/ml) and from 0 to 847.5 μg/ml (mean: 7.0 μg/ml, median: 2.5 μg/ml), respectively. For SAA positivity, about a 2-fold increased risk was observed for total gastric cancer and non-cardia cancer; the adjusted ORs (95% CIs) were 1.93 (1.22–3.07, $P = 0.005$) and 2.13 (1.14–3.98, $P = 0.02$), respectively (Table IV). Among adjusted covariates, atrophy as well as *H.pylori* infection contributed most of the elevation of risk. Among non-cardia cancers, no difference was observed by histologic type. The largest OR was demonstrated for cardia cancers, although it failed to reach the level of statistical significance; the adjusted OR equaled 3.84 (95% CI: 0.82–17.99, $P = 0.09$). When results for SAA status were shown separately for men and women, there was no material difference; the adjusted ORs for developing total gastric cancer were 1.95 and 2.15 for men and women, respectively. The adjusted OR for cardia cancer among women could not be calculated because of the small sample size; therefore, all analyses were conducted for men and women combined. No apparent association was observed when SAA

Table III. ORs and 95% CIs of CRP positivity (CRP > 0.18 mg/dl) for development of gastric cancer by tumor subsite and histologic type

	No. of CRP-positive cases/controls	Crude OR (95% CI)	*P*-value	Adjusted OR (95% CI)[a]	*P*-value
All (494 pairs)	62/47	1.36 (0.91–2.02)	0.13	1.90 (1.19–3.02)	0.007
Cardia (39 pairs)	7/2	3.50 (0.73–16.85)	0.12	3.14 (0.51–19.39)	0.22
Non-cardia (355 pairs)	44/33	1.36 (0.85–2.16)	0.20	2.18 (1.24–3.84)	0.007
Differentiated type (232 pairs)	30/23	1.32 (0.76–2.29)	0.33	1.77 (0.89–3.52)	0.10
Undifferentiated type (107 pairs)	9/8	1.14 (0.41–3.15)	0.80	2.01 (0.53–7.62)	0.30

[a]Cardia cancers, adjusted for cigarette smoking, BMI, family history of gastric cancer, history of infectious or inflammatory disease, current drug use of analgesics and salt intake. All gastric cancers, all non-cardia cancers, differentiated-type non-cardia cancer and undifferentiated-type non-cardia cancer, further adjusted for *Helicobacter pylori* infection, CagA positivity and atrophy.

Table IV. ORs and 95% CIs of SAA positivity (SAA > 8 μg/ml) for development of gastric cancer by tumor subsite and histologic type

	No. of SAA-positive cases/controls	Crude OR (95% CI)	*P*-value	Adjusted OR (95% CI)[a]	*P*-value
All (494 pairs)	75/63	1.26 (0.86–1.86)	0.24	1.93 (1.22–3.07)	0.005
Cardia (39 pairs)	11/5	3.00 (0.81–11.08)	0.10	3.84 (0.82–17.99)	0.09
Non-cardia (355 pairs)	45/39	1.21 (0.74–2.00)	0.45	2.13 (1.14–3.98)	0.02
Differentiated type (232 pairs)	27/25	1.11 (0.59–2.06)	0.75	1.73 (0.81–3.72)	0.16
Undifferentiated type (107 pairs)	12/11	1.14 (0.41–3.15)	0.80	1.80 (0.41–7.92)	0.44

[a]Cardia cancers, adjusted for cigarette smoking, alcohol consumption, BMI, family history of gastric cancer, history of infectious or inflammatory disease, current drug use of analgesics and salt intake. All gastric cancers, all non-cardia cancers, differentiated-type non-cardia cancer and undifferentiated-type non-cardia cancer, further adjusted for *Helicobacter pylori* infection, CagA positivity and atrophy.

level was divided into quartiles (<1.3, 1.3–2.5, 2.5–5.1 and ≥5.1 μg/ml). Compared with the lowest (reference) group, the adjusted ORs (95% CIs) for development of gastric cancer for the second, the third and the highest group were 0.81 (0.53–1.24), 1.06 (0.70–1.61) and 1.19 (0.77–1.85), respectively ($P = 0.20$). When SAA was treated as a continuous measure, the adjusted OR for development of gastric cancer was 1.00 (0.995–1.00) for 1 mg/dl increase of log-transformed SAA. Similar to the analysis of CRP, non-linear continuous models did not reveal any evidence of dose response.

Because of the high correlation between CRP and SAA, we included only the values for the marker being analyzed (Tables III and IV). When CRP and SAA were included in the model simultaneously, the OR was attenuated and was no longer significant for CRP, but was still significant for SAA (data not shown). This may not contradict previous reports that suggest overlapping of the roles of the two markers and delineation of SAA as the more sensitive parameter (13,14).

The observed association did not differ for stratification by smoking status (never/past + current) for SAA; however, for CRP, the association was clearer among never smokers [2.50 (1.13–5.53)] compared with past and current smokers [1.15 (0.56–2.33)]. Using the World Health Organization category to adjust BMI did not alter the results essentially. When the interactions between each covariate in the model and CRP and SAA status were tested, no significant interaction was observed.

When all analyses were repeated in only those who were *H.pylori* positive (seropositive for IgG antibody and/or CagA), the associations were slightly attenuated, although they did not differ essentially; the adjusted ORs (95% CIs) for developing total gastric cancer were 1.72 (1.07–2.78, $P = 0.03$) for CRP-positive status and 1.82 (1.13–2.94, $P = 0.01$) for SAA-positive status, respectively.

Discussion

In this study, the overall distributions of CRP and SAA were not apparently associated with the development of gastric cancer. However, when subjects were divided on the basis of dichotomous categorization of positive versus negative, an increased risk was observed for positive subjects. The association was statistically significant even after adjustment for *H.pylori* infection, CagA status, atrophy and lifestyle factors. Elevated levels of CRP and SAA reflect a generalized host reaction that is either localized or systematic with regard to the initial event. Mechanisms of inflammation-associated tumor development are well described. These include stimulation of cellular proliferation (e.g. in cellular proto-oncogenes, DNA and cellular repair), inhibition of apoptosis, cellular adhesion, stimulation of angiogenesis and cellular transformation (1). In our data set, under the conditions that most subjects were infected with *H.pylori*, only those who reacted strongly to inflammation or had a high host immune response, as reflected by extremely elevated plasma levels of CRP and SAA, showed an elevated risk of developing the malignancy. The proportions of those who were categorized as positive were small; therefore, the findings should be interpreted with caution. However, this may be one of the explanations for why only a small proportion of *H.pylori*-infected subjects develop clinical consequences. CRP and SAA were useful markers to detect these high-risk groups.

Several clinical studies have shown that, compared with controls, gastric cancer patients have elevated CRP levels (24–26). Previous studies have even revealed that CRP has an impact on gastric cancer prognosis (24,27). It has been observed in previous studies that the SAA level increases in patients with stomach, lung, renal, colorectal, breast and other forms of cancers (28–35). With regard to gastric cancer, Chan *et al.* (28) demonstrated that patients with gastric cancer have higher SAA concentrations than do patients with gastric ulcers and healthy subjects and that levels of SAA correlate with tumor status, prognosis and recurrence. In our study, the average duration between blood donation and cancer diagnosis among cases was 5.4 years. When subjects who developed gastric cancer within 2 years of blood donation and their matched controls were excluded, the observed associations were strengthened; the adjusted ORs (95% CIs) for the association between development of gastric cancer and CRP and SAA positivity were 2.25 (1.31–3.85, $P = 0.003$) and 2.29 (1.32–3.95, $P = 0.003$), respectively. Furthermore, when subjects were stratified by the median duration between blood donation and

diagnosis (5.12 years), the adjusted ORs (95% CIs) for the association between development of gastric cancer CRP and SAA positivity within 5.12 years were 1.38 (0.69–2.73, $P = 0.36$) and 1.59 (0.82–3.09, $P = 0.17$), respectively. The values for CRP and SAA diagnosed after 5.12 years were 2.42 (1.23–4.77, $P = 0.01$) and 2.25 (1.12–4.52, $P = 0.02$), respectively. Therefore, our findings cannot be explained by the effect of preclinical samples among cases. Rather, our findings suggest that CRP and SAA may be useful markers for predicting the malignancy.

In our study, *H.pylori* seropositivity, *H.pylori* titer and atrophy were not distributed differently according to CRP status. Surprisingly, *H.pylori* seropositivity and atrophy were more frequent, and higher *H.pylori* titer was observed among SAA-negative subjects than among SAA-positive subjects. When the values were compared on the basis of tumor location, CRP did not show any difference; mean value (SE) was 0.09 (0.16) for cardia and 0.15 (0.05) for non-cardia cancer, respectively ($P = 0.75$). The value for SAA was 6.77 (1.87) for cardia, which was higher than that for non-cardia, 5.12 (0.62) ($P = 0.03$). High SAA level with an upper tumor site compared with a middle or a lower site was also observed by Chan *et al.* (28). Furthermore, the largest OR was observed for cardia cancer for both CRP and SAA. It is well known that *H.pylori* infection is related to non-cardia gastric cancer. As the majority of our subjects were infected with *H.pylori*, we were unable to show the results among *H.pylori*-seronegative subjects. Therefore, we cannot clarify whether the observed phenomenon was independent of *H.pylori*. We can state only that the observed elevated risk of gastric cancer with high levels of CRP and SAA is probably a phenomenon that cannot be totally explained by *H.pylori*; this conclusion is in line with that of previous studies (26,36). Comparing 153 preoperative gastric cancer patients with 19 healthy subjects, Tsavaris *et al.* (26) observed high serum levels of CRP, ceruloplasmin and a1-acid glycoprotein in cancer patients; however, among cancer patients, CRP level did not differ by status of *H.pylori* infection. Also, Delanghe *et al.* (36) showed that neither SAA nor other acute-phase proteins, including CRP, correlated with *Chlamydia pneumoniae* IgG, *H.pylori* IgG and IgA and cytomegalovirus IgG. On the other hand, the reason for the large OR observed in the cardia for both CRP and SAA positivity is unknown. One recent study reported that plasma CRP levels were associated with high BMI and other indicators of obesity (37). On the other hand, some studies, but not all, have proposed that elevated body weight may increase the risk of gastroesophageal reflux, which has been associated with adenocarcinomas of the gastroesophageal junction (38). Therefore, it is possible that elevated CRP and SAA were strongly associated with cardia cancer because of BMI status. However, in our data set, BMI did not differ by either CRP status or SAA status. The observed high OR in cardia cancer may be due to factors other than BMI or may be a mere chance finding.

On the basis of self-reported information, we adjusted for any condition that might alter the plasma levels of CRP or SAA. When these subjects were deleted (61 pairs; corresponds to 12% of total subjects), the overall findings did not change essentially, except when CRP values were divided into quartiles; the P for trend then became not significant ($P = 0.44$). Alternatively, when subjects with an extremely high level of CRP (>0.5 mg/dl) or SAA (>16.5 μg/ml) were excluded (55 pairs; corresponds to 11% of total subjects), the observed ORs became slightly higher, although the overall findings did not change essentially. To ensure the generalizability of findings and statistical power, we retained these subjects in the analyses.

Our study has several limitations. First, among 97 644 eligible subjects of the JPHC Study cohort, 36 745 (38%) men and women participated in the survey and provided blood samples. As reported previously, compared with non-participants, participants in the health checkup survey, especially women, had a different socioeconomic status and a favorable lifestyle profile, such as less smoking and alcohol consumption, greater participation in physical exercise and greater consumption of fruits or green vegetables (39). These findings mean that caution is needed in generalizing or interpreting the results in this report. Second, because of the relatively small sample size, further studies are needed to test our findings in analyses conducted by tumor location and histologic subtype.

The advantage of this study is its population-based prospective design and analysis of prediagnosed blood samples. Also, detailed information including *H.pylori* infection, CagA status, atrophy and environmental factors contributed to the detection of the relationships independent of these factors. Other strengths include negligible loss to follow-up and the satisfactory quality of our cancer registry system during the study period.

In conclusion, the overall distribution of CRP and SAA was not apparently associated with the development of gastric cancer. However, it was suggested that those who react strongly to inflammation or who have high host immune response, as reflected by extremely elevated plasma levels of CRP and SAA, were at high risk to develop gastric cancer.

Funding

Ministry of Health, Labour and Welfare of Japan [Grant-in-Aid for Cancer Research (19 shi-2); Third Term Comprehensive 10-year Strategy for Cancer Control (H21-Sanjigan-Ippan-003)]; Ministry of Education, Culture, Sports, Science, and Technology of Japan and Japan Society for the Promotion of Science [Grants-in-Aid for Scientific Research for Young Scientists (A), 19689014].

Acknowledgements

We thank all staff members in each study area for their painstaking efforts to conduct the baseline survey and follow-up.

Conflict of Interest Statement: None declared.

References

1. Macarthur,M. *et al.* (2004) Inflammation and cancer II. Role of chronic inflammation and cytokine gene polymorphisms in the pathogenesis of gastrointestinal malignancy. *Am. J. Physiol. Gastrointest. Liver Physiol.*, **286**, G515–G520.
2. Yamaoka,Y. *et al.* (2001) Relation between cytokines and *Helicobacter pylori* in gastric cancer. *Helicobacter*, **6**, 116–124.
3. Wu,C.W. *et al.* (1996) Serum interleukin-6 levels reflect disease status of gastric cancer. *Am. J. Gastroenterol.*, **91**, 1417–1422.
4. Kabir,S. *et al.* (1995) Serum levels of interleukin-1, interleukin-6 and tumor necrosis factor-alpha in patients with gastric carcinoma. *Cancer Lett.*, **95**, 207–212.
5. Rad,R. *et al.* (2004) Cytokine gene polymorphisms influence mucosa cytokine expression, gastric inflammation, and host specific colonization during *Helicobacter pylori* infection. *Gut.*, **53**, 1082–1089.
6. El-Omar,E.M. *et al.* (2000) Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature*, **404**, 398–402.
7. Togawa,S. *et al.* (2005) Interleukin-2 gene polymorphisms associated with increased risk of gastric atrophy from *Helicobacter pylori* infection. *Helicobacter*, **10**, 172–178.
8. Taguchi,A. *et al.* (2005) Interleukin-8 promoter polymorphism increases the risk of atrophic gastritis and gastric cancer in Japan. *Cancer Epidemiol. Biomakers Prev.*, **14**, 2487–2493.
9. Mackiewicz,A. *et al.* (1991) Effects of cytokine combinations on acute phase protein production in two human hepatoma cell lines. *J. Immunol.*, **146**, 3032–3037.
10. Gabay,C. *et al.* (1999) Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med.*, **340**, 448–454.
11. Mendall,M.A. *et al.* (1997) Relation of serum cytokine concentrations to cardiovascular risk factors and coronary heart disease. *Heart*, **78**, 273–277.
12. O'Hanlon,D.M. *et al.* (2002) The acute phase response in breast cancer. *Anticancer Res.*, **22**, 1289–1294.
13. Maury,C.P.J. (1985) Comprehensive study of serum amyloid A protein and C-reactive protein in disease. *Clin. Sci.*, **68**, 233–238.
14. Marhaug,G. *et al.* (1986) Serum amyloid A protein in acute myocardial infarction. *Acta Med. Scand.*, **220**, 303–306.
15. Watanabe,S. *et al.* (2001) Study design and organization of the JPHC Study. *J. Epidemiol.*, **11** (Suppl): S3–S7.
16. Tsugane,S. *et al.* (2001) Dietary habits among the JPHC Study participants at baseline survey. *J. Epiemiol.*, **11** (Suppl): S44–S56.

17. World Health Organization. (1990) *International Classification of Diseases for Oncology*, 2nd ed. World Health Organization, Geneva.
18. Japanese Research Society for Gastric Cancer. (1993) *The General Rules for Gastric Cancer Study*, 12th ed. Kanehara, Tokyo.
19. Lauren,P. (1965) The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. *APMIS*, **64**, 31–49.
20. Kobayashi,T. *et al.* (2003) The basic efficiency of LZ test 'Eiken' CRP-HG. *Iryou to Kensakiki, Shinyaku*, **26**, 497–504. (in Japanese).
21. Kousaka,T. *et al.* (1994) Clinical evaluation of serum amyloid A (SAA) by latex agglutination nephelometric immunoassay. *Igaku to Yakugaku*, **31**, 1191–1210. (in Japanese).
22. Miki,K. (1998) *Pepsinogen Method*. Igakushoin, Tokyo (in Japanese).
23. Kabuto,M. *et al.* (1993) Correlation between atrophic gastritis prevalence and gastric cancer mortality among middle-aged men in 5 areas in Japan. *J. Epidemiol.*, **3**, 35–39.
24. Wu,C.W. *et al.* (1988) Alterations of humoral immunity in patients with gastric cancer. *Asian Pac. J. Allergy Immunol.*, **6**, 7–10.
25. Ilhan,N. *et al.* (2004) C-reactive protein, procalcitonin, interleukin-6, vascular endothelial growth factor and oxidative metabolites in diagnosis of infection and staging in patients with gastric cancer. *World J. Gastroenterol.*, **10**, 1115–1120.
26. Tsavaris,N. *et al.* (2005) Retinol-binding protein, acute phase reactants and *Helicobacter pylori* infections in patients with gastric adenocarcinoma. *World J. Gastroenterol.*, **11**, 7174–7178.
27. Rashid,S.A. *et al.* (1982) Plasma protein profiles and prognosis in gastric cancer. *Br. J. Cancer*, **45**, 390–394.
28. Chan,D.-C. *et al.* (2006) Evaluation of serum amyloid A as a biomarker for gastric cancer. *Ann. Surg. Oncol.*, **14**, 84–93.
29. Biran,H. *et al.* (1986) Serum amyloid A (SAA) variations in patients with cancer: correlation with disease activity, stage, primary site, and prognosis. *J. Clin. Pathol.*, **39**, 794–797.
30. Kimura,M. *et al.* (2001) Significance of serum amyloid A on the prognosis in patients with renal cell carcinoma. *Cancer*, **92**, 2072–2075.
31. Glojnaric,I. *et al.* (2001) Serum amyloid A protein (SAA) in colorectal carcinoma. *Clin. Chem. Lab. Med.*, **39**, 129–133.
32. Kaneti,J. *et al.* (1984) Importance of serum amyloid A (SAA) level in monitoring disease activity and response to therapy in patients with prostate cancer. *Urol. Res.*, **12**, 239–241.
33. Cho,W.C.S. *et al.* (2004) Identification of serum amyloid a protein as a potentially useful biomarker to monitor relapse of nasopharyngeal cancer by serum proteomic profiling. *Clin. Cancer Res.*, **10**, 43–52.
34. O'Hanlon,D.M. *et al.* (2002) The acute phase response in breast carcinoma. *Anticancer Res.*, **22**, 1289–1293.
35. Rosenthal,C.J. *et al.* (1979) Serum amyloid A to monitor cancer dissemination. *Ann. Intern. Med.*, **91**, 383–390.
36. Delanghe,J.R. *et al.* (2002) Discriminative value of serum amyloid A and other acute-phase proteins for coronary heart disease. *Atherosclerosis*, **160**, 471–476.
37. Timpson,N.J. *et al.* (2005) C-reactive protein and its role in metabolic syndrome: Mendelian randomization study. *Lancet*, **366**, 1954–1959.
38. Yang,P. *et al.* (2009) Overweight, obesity and gastric cancer risk: results from a meta-analysis of cohort studies. *Eur. J. Cancer*, **45**, 2867–2873.
39. Iwasaki,M. *et al.* (2003) Background characteristics of basic health examination participants: the JPHC Study baseline survey. *J. Epidemiol.*, **13**, 216–225.

Received June 1, 2010; revised January 10, 2010; accepted January 10, 2010

Appendix

Members of the JPHC Study Group: S. Tsugane (principal investigator), M. Inoue, T. Sobue and T. Hanaoka, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo; J. Ogata, S. Baba, T. Mannami, A. Okayama and Y. Kokubo, National Cardiovascular Center, Suita; K. Miyakawa, F. Saito, A. Koizumi, Y. Sano, I. Hashimoto, T. Ikuta and Y. Tanaba, Iwate Prefectural Ninohe Public Health Center, Ninohe; Y. Miyajima, N. Suzuki, S. Nagasawa, Y. Furusugi and N. Nagai, Akita Prefectural Yokote Public Health Center, Yokote; H. Sanada, Y. Hatayama, F. Kobayashi, H. Uchino, Y. Shirai, T. Kondo, R. Sasaki, Y. Watanabe, Y. Miyagawa, Y. Kobayashi and M. Machida, Nagano Prefectural Saku Public Health Center, Saku; Y. Kishimoto, E. Takara, T. Fukuyama, M. Kinjo, M. Irei and H. Sakiyama, Okinawa Prefectural Chubu Public Health Center, Okinawa; K. Imoto, H. Yazawa, T. Seo, A. Seiko, F. Ito, F. Shoji and R. Saito, Katsushika Public Health Center, Tokyo; A. Murata, K. Minato, K. Motegi and T. Fujieda, Ibaraki Prefectural Mito Public Health Center, Mito; T. Abe, M. Katagiri, M. Suzuki and K. Matsui, Niigata Prefectural Kashiwazaki and Nagaoka Public Health Center, Kashiwazaki and Nagaoka; M. Doi, A. Terao, Y. Ishikawa and T. Tagami, Kochi Prefectural Chuo-higashi Public Health Center, Tosayamada; H. Doi, M. Urata, N. Okamoto, F. Ide and H. Sueta, Nagasaki Prefectural Kamigoto Public Health Center, Arikawa; H. Sakiyama, N. Onga, H. Takaesu and M. Uehara, Okinawa Prefectural Miyako Public Health Center, Hirara; F. Horii, I. Asano, H. Yamaguchi, K. Aoki, S. Maruyama, M. Ichii and M. Takano, Osaka Prefectural Suita Public Health Center, Suita; S. Matsushima and S. Natsukawa, Saku General Hospital, Usuda; M. Akabane, Tokyo University of Agriculture, Tokyo; M. Konishi, K. Okada and I. Saito, Ehime University, Toon; H. Iso, Osaka University, Suita; Y. Honda, K. Yamagishi, S. Sakurai and N. Tsuchiya, Tsukuba University, Tsukuba; H. Sugimura, Hamamatsu University, Hamamatsu; Y. Tsubono, Tohoku University, Sendai; M. Kabuto, National Institute for Environmental Studies, Tsukuba; S. Tominaga, Aichi Cancer Center Research Institute, Nagoya; M. Iida, W. Ajiki and A. Ioka, Osaka Medical Center for Cancer and Cardiovascular Disease, Osaka; S. Sato, Osaka Medical Center for Health Science and Promotion, Osaka; N. Yasuda, Kochi University, Nankoku; K. Nakamura, Niigata University, Niigata; S. Kono, Kyushu University, Fukuoka; K. Suzuki, Research Institute for Brain and Blood Vessels Akita, Akita; Y. Takashima and M. Yoshida, Kyorin University, Mitaka; E. Maruyama, Kobe University, Kobe; M. Yamaguchi, Y. Matsumura, S. Sasaki and S. Watanabe, National Institute of Health and Nutrition, Tokyo; T. Kadowaki, Tokyo University, Tokyo; M. Noda and T. Mizoue, International Medical Center of Japan, Tokyo; Y. Kawaguchi, Tokyo Medical and Dental University, Tokyo; and H. Shimizu, Sakihae Institute, Gifu.

Cancer Epidemiology 34 (2010) 534–541

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.canceepidemiology.net

10-Year risk of colorectal cancer: Development and validation of a prediction model in middle-aged Japanese men

Enbo Ma, Shizuka Sasazuki *, Motoki Iwasaki, Norie Sawada, Manami Inoue

Shoichiro Tsugane for the Japan Public Health Center-based Prospective Study Group

Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan

ARTICLE INFO

Article history:
Accepted 30 April 2010
Available online 31 May 2010

Keywords:
Prospective study
Epidemiologic method
Risk prediction
Colorectal cancer
Validation
Japanese men

ABSTRACT

Background: To estimate an individual's probability of developing colorectal cancer (CRC) may aid health professionals and individuals in improving lifestyle behaviors or deciding the screening regimens. As fewer studies on cancer risk prediction were seen so far, we initially developed an assessment tool with synthesizing key information from a variety of CRC risk factors through a large population-based cohort study. *Method:* The prediction model was derived from 28,115 men in the Japan Public Health Center-based (JPHC) Prospective Study Cohort II (follow-up: 1993–2005), with risk factors selected by Cox proportion hazard regression. 18,256 men in the JPHC Study Cohort I (follow-up: 1995–2005) were used to evaluate the model's performance. *Results:* 543 and 398 CRCs were diagnosed during the follow-up period in Cohorts II and I, respectively. The prediction model, including age, BMI, alcohol consumption, smoking status, and the daily physical activity level, showed modest discrimination ability for CRC ($C = 0.70$; 95% confidential interval, 0.68–0.72) in Cohort II and well calibrated in Cohort I (Hosmer–Lemeshow $\chi^2 = 14.2$, $P = 0.08$). *Conclusion:* The 10-year CRC risk prediction model may be used to estimate CRC risk in Japanese men. It may also play a role in the promotion of CRC prevention strategies.

1. Introduction

Colorectal cancer (CRC) was the second most commonly diagnosed cancer in the Japanese population in 2002 [1,2]. Approximately 11% of total cancer deaths in men and 14% in women were from CRCs in 2005 [2]. The high morbidity and mortality noted in the Japanese population were similar to those in North American and European counties [3].

Some risk factors for CRC were documented in the revised expert report from the World Cancer Research Fund, including physical activity, alcohol consumption, body and abdominal fatness, and consumption of vegetables and foods containing fiber [4]. A recent meta-analysis confirmed that smoking was significantly associated with CRC incidence and mortality [5]. In epidemiologic studies of the Japanese population, the risk factors of physical activity [6,7], alcohol consumption [8,9], smoking habit [8,9], and body mass index (BMI) [9,10] were consistently identified, whereas consumption of vegetables [11] and foods containing fiber [12] were not. Systematic reviews of large studies in Japan also verified the findings for alcohol consumption [13] and smoking habit [14]. In the Japanese population, however, these risk factors were more prevalent in men than in women, and little evidence of modifying CRC risk by reproductive factors has been found among Japanese women [15,16]. Nevertheless, most of these established risk factors for CRC are modifiable, and their improvement has been incorporated into primary cancer prevention strategies in Japan [17].

Given the high incidence of CRC and its significant cost to society, it is critical to reduce the identified risk factors in order to prevent CRC in a population. An individual's risk probability of developing CRC could be estimated by using information on established factors, which would aid physicians and individuals in improving lifestyle behavior and/or deciding on screening regimens for CRC prevention [17–19]. Moreover, from the public health point of view, risk prediction tools could also be used to effectively disseminate information on cancer prevention.

Several studies estimated the absolute risk probability of developing CRC, although they were based on case–control study [18], expert opinion [20], or specific populations [21,22]. In this paper, we present a CRC risk prediction model in Japanese men, derived and validated by two large cohorts from the Japan Public Health Center-based (JPHC) Prospective Study. We also present a simplified score model that can be easily used to estimate an individual's absolute CRC risk based on lifestyle information.

* Corresponding author. Tel.: +81 3 3542 2511x3378; fax: +81 3 3547 8578.
E-mail address: ssasazuk@ncc.go.jp (S. Sasazuki).

1877-7821/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.canep.2010.04.021

2. Materials and methods

2.1. Study participants

In the JPHC Study, Cohort I, with participants aged 40–59 years, was launched in 1990 and Cohort II, with participants aged 40–69 years, was added in 1993. A total of 48,448 men were initially identified in 11 public health center-based (PHC) areas throughout Japan. The details of the study design and baseline response have been described elsewhere [23,24]. The study was approved by the Institute Review Board of the National Cancer Center, Tokyo, Japan.

The baseline survey for Cohort II had more comprehensive data on physical activity and the food frequency questionnaire (FFQ) (52 food items) than those and the FFQ (44 food items) for Cohort I. In the 5-year follow-up survey, all investigations including the FFQ (138 food items) were the same for both cohorts. Considering the inconsistency of questionnaires and follow-up periods of the two cohorts, in the present study we used the baseline survey of Cohort II men to derive the risk prediction model of CRC and the 5-year follow-up survey of Cohort I men to validate the model.

Participants who reported a history of cancer or cardiovascular disease, were diagnosed with cancers, or were censored before the start of the follow-up survey were excluded, leaving 28,115 eligible subjects for model derivation in Cohort II and 18,256 for model validation in Cohort I.

2.2. Risk factor measurements

Self-administered questionnaires contained items on demographic characteristics, medical history, smoking habit, alcohol consumption, physical activity, occupation, and other factors, as well as diets by validated FFQs [25,26].

BMI was calculated as weight in kilograms divided by the square of height in meters. Physical activity levels, measured by metabolic equivalent (MET) hours per day, were estimated by multiplying the reported time spent at each activity per day by its assigned MET intensity: heavy physical work or strenuous exercise (4.5), walking or standing (2.0), sedentary (1.5), and sleep or others (0.9) [6,27]. Daily physical activity level was the sum of MET-hour scores across all activities.

Smoking habit was grouped into never, former, and current smokers. Alcohol consumption was categorized into four groups (never, occasional, regular <300 g/week, and regular ≥300 g/week), in which regular drinkers were categorized by multiplying the frequency per week by the usual daily amount of alcohol consumed [8].

Daily food intake was calculated by multiplying the frequency by standard portion size and relative size for each food item in the FFQ. Daily intake of nutrients was calculated using the 5th revised edition of the Standard Tables of Food Composition in Japan [28].

2.3. Follow-up and case assessment

Participants were followed until 31 December 2005. Residence status, movement of households, and survival were confirmed annually using the residential registers. Information on the cause of death was obtained by examining the death certificates provided by the Ministry of Health, Labour, and Welfare. The occurrence of cancer was identified by active patient notification through the major local hospitals in the study areas and data linkage with population-based cancer registries. The site and histology of each cancer were coded using the International Classification of Diseases for Oncology, 3rd edition (ICD-O-3), with C18–C20 for CRC, C180–C189 for colon cancer, and C199 and C209 for rectal cancer.

2.4. Statistical analysis

Person-years of follow-up were counted from the date of survey response (1993 for Cohort II and 1995 for Cohort I) until the date of CRC diagnosis, the date of moving out of a study area, the date of death, or the end of 2005, whichever came first. Persons lost to follow-up were censored on the last confirmed date of their presence in the study area. Extreme values of height (<100 or >199 cm), weight (<20 kg), and BMI (<14 or >40 kg/m^2) were removed from this analysis. Nutrient intakes were categorized into tertiles for all study participants, with the lower tertile as the reference.

2.4.1. Prediction model derived by JPHC Cohort II

Cox proportional hazards models were derived after testing for the assumptions underlying its use. Then the model of predictive risk of developing CRC was fitted, in which the average survival rates at follow-up time points were estimated by baseline hazard function with mean values of potential predictors. Hazard ratios (HR) and 95% confidential interval (CI) of each risk factor were also estimated. Based on the previous publications in Japanese populations and age-adjusted univariate analysis performed for available variables in this study (including more than 30 food items and nutrients), the potential predictors were applied for building the full multivariate model, which including age, BMI, daily physical activity, alcohol consumption, smoking habit, family history of CRC, and diabetes diagnosed, and interested interaction terms with biological plausibility between alcohol and smoking, and physical activity and BMI. PHC areas were treated as strata in the analysis; assessment of likely shrinkage (over-fitting) was evaluated for the reduced models by $[LR-(p-q)-q]/[LR-(p-q)]$, where LR denotes the likelihood ratio χ^2, and p and q denote the regression degrees of freedom for the full model and for a reduced model, respectively [29]. Non-linear relationships (transformations) of age, BMI, or daily physical activity were tested by using multiple fractional polynomial method of two degree [30,31], however, none of which had been statistically significant for leaving in the model.

For each risk factor, the regression coefficients of two cohorts were compared by a 2-tailed Z statistics, $Z=(\beta_{[d]}-\beta_{[v]})/SE$, where $\beta_{[d]}$ and $\beta_{[v]}$ are the regression coefficients of Cohort II and Cohort I, respectively, and SE is the standard error of the difference in the coefficients, calculated as $\sqrt{(SE^2_{\beta_{[d]}}\pm SE^2_{\beta_{[v]}})}$ [32]. The Z statistic was used to test the difference in HR of each risk factor/category between the two cohorts [32]. The individual risk of CRC was estimated based on the baseline hazard function of the Cox regression model derived from Cohort II, which method was same as one developed in Framingham heart study [33], where $P=1-S(t)^{\exp(f[x,M])}$ and $f(x,M)=\beta 1(x1-M1)+\ldots+\beta j(xj-Mj)$. $\beta 1,\ldots,\beta j$ are the regression coefficients, $x1,\ldots,xj$ represent an individual's risk factors, $M1,\ldots,Mj$ are the mean values of the risk factors in the cohort (for category variables, $x1,\ldots,xj$ are the dichotomous value of the created dummy variable for each category, entering 1 if the individual's value fits that certain category and 0 otherwise, and $M1,\ldots,Mj$ are the proportion of the certain category of the variable in the cohort), and $S(t)$ is the average survival rate at time t of subjects with the mean values of the risk factors used in the Cox model. This procedure performed a better validity than prepared by Ederer method [34]. The predicted 10-year risk of CRC, therefore, was estimated by the baseline hazard function of Cohort II with mean values of each predictor at the 10-year follow-up time.

2.4.2. Prediction model validated by JPHC Cohort I

Discrimination, the ability of a predictive model to separate those who experience an event from those who do not, was

Table 1
Full and reduced predicative models for estimation of developing colorectal cancer events in Cohort II men, Japan Public Health Center-based Prospective Study, 1993–2005.

Variables retained	Full model			Reduced 1[a]			Reduced 2[b]		
	β	S.E.(β)	*P*-Value	β	S.E.(β)	*P*-Value	β	S.E.(β)	*P*-Value
CRC[c]									
Age, year	0.079	0.006	<0001	0.080	0.006	<0001	0.080	0.006	<0001
BMI, kg/m^2	0.001	0.061	0.98	0.047	0.016	<0.01	0.047	0.016	<0.01
Physical activity, MET-h/d	−0.055	0.049	0.27	−0.019	0.006	<0.01	−0.019	0.006	0.01
Family history of CRC (yes)	−0.085	0.382	0.82	−0.087	0.382	0.82	–	–	–
Diabetes (yes)	0.103	0.160	0.52	0.095	0.160	0.55	–	–	–
Alcohol consumption[d]									
Never	0.052	0.244	0.83	−0.163	0.210	0.44	−0.163	0.210	0.44
Regular (<300 g/w)	0.393	0.230	0.09	0.359	0.192	0.06	0.358	0.192	0.06
Regular (≧300 g/w)	0.584	0.273	0.03	0.657	0.195	0.001	0.659	0.195	0.001
Smoking									
Former	−0.165	0.196	0.40	0.070	0.133	0.60	0.071	0.133	0.59
Current	−0.225	0.330	0.50	0.237	0.119	0.05	0.239	0.119	0.04
Smoking × alcohol	0.078	0.056	0.17	–	–	–	–	–	–
BMI × physical activity	0.002	0.002	0.46	–	–	–	–	–	–
d.f.	12			10			8		
Likelihood ratio x^2	239.8			237.3			241.2		
Shrinkage	–			0.96			0.97		
C-Index	0.703			0.699			0.699		
Colon cancer									
Age, year	0.084	0.008	<0001	0.085	0.008	<0001	0.085	0.008	<0001
BMI, kg/m^2	0.037	0.079	0.64	0.048	0.021	0.02	0.049	0.021	0.02
Physical activity, MET-h/d	−0.028	0.063	0.66	−0.019	0.008	0.02	−0.020	0.008	0.01
Family history of CRC (yes)	0.438	0.384	0.25	0.437	0.384	0.26	–	–	–
Diabetes (yes)	0.330	0.188	0.08	0.323	0.188	0.09	–	–	–
Alcohol consumption[d]									
Never	0.077	0.323	0.81	−0.133	0.276	0.63	−0.140	0.276	0.61
Regular (<300 g/w)	0.493	0.305	0.11	0.431	0.253	0.09	0.419	0.254	0.10
Regular (≧300 g/w)	0.651	0.363	0.07	0.657	0.257	0.01	0.655	0.258	0.01
Smoking									
Former	−0.006	0.258	0.98	0.180	0.173	0.30	0.186	0.173	0.28
Current	−0.012	0.433	0.98	0.341	0.157	0.03	0.347	0.157	0.03
Smoking × alcohol	0.057	0.073	0.44	–	–	–	–	–	–
BMI × physical activity	0.000	0.003	0.90	–	–	–	–	–	–
d.f.	12			10			8		
Likelihood ratio x^2	165.7			165.1			166.0		
Shrinkage	–			0.94			0.95		
C-Index	0.710			0.710			0.708		
Rectal cancer									
Age, year	0.072	0.009	<0001	0.071	0.009	<0001	0.067	0.009	<0001
BMI, kg/m^2	−0.054	0.098	0.58	0.033	0.025	0.19	–	–	–
Physical activity, MET-h/d	−0.097	0.078	0.22	−0.018	0.010	0.07	−0.020	0.008	0.02
Diabetes (yes)	−0.357	0.311	0.25	−0.078	0.240	0.75	–	–	–
Alcohol consumption[d]									
Never	0.027	0.374	0.94	−0.401	0.291	0.17	−0.094	0.361	0.80
Regular (<300 g/w)	0.261	0.349	0.45	0.083	0.259	0.75	0.365	0.335	0.28
Regular (≧300 g/w)	−0.536	0.514	0.30	0.488	0.268	0.07	0.745	0.281	0.01
Smoking									
Former	−0.3%	0.305	0.19	0.088	0.181	0.63	–	–	–
Current	0.504	0.415	0.22	0.088	0.181	0.63	–	–	–
Smoking × alcohol	0.109	0.087	0.21	–	–	–	–	–	–
BMI × physical activity	0.003	0.003	0.31	–	–	–	–	–	–
d.f.	11			9			5		
Likelihood ratio χ^2	82.9			80.0			75.7		
Shrinkage	–			0.89			0.94		
C-Index	0.698			0.678			0.678		

[a] Removed interactions.
[b] Further removed family history and diabetes diagnosed for CRC and colon cancer; diabetes diagnosed, BMI, and smoking habit for rectal cancer.
[c] CRC, colorectal cancer; MET, metabolic equivalent.
[d] Occasional alcohol consumption was as the reference.

assessed using the *C* statistic, the area under the receiver operating characteristic curve [32]. The overall *C* statistics and its 95% CIs were calculated by logistic regressions. Calibration is another measure of performance of a prediction model that tests how closely predicted outcomes agree with actual outcomes [32,35]. The calibration was conducted in Cohort I, using the β coefficients, the mean of each risk factor, and the average survival rate at 10-year from the original Cohort II. Participants in Cohort I were divided into 10 deciles of individual predicted risk, and in each decile the expected events were the sum of individual predicted

Table 2
Characteristics of risk factors, person-years of follow-up, and colorectal cancer events in men, Japan Public Health Center-based Prospective Study, 1993–2005[a].

Risk factor	Cohort II[b]						Cohort I[c]					
	Participants, mean (SD), %	No. of participants	Person-years of follow-up	No. of events			Participants, mean (SD), %	No. of participants	Person-years of follow-up	No. of events		
				CRC	Colon	Rectum				CRC	Colon	Rectum
Age, year	52.9(8.8)	28,115	310,059	543	329	214	54.7 (6.0)	18,256	184,496	389	239	150
BMI, kg/m^2	23.4 (2.9)	28,115	310,059	543	329	214	23.6 (2.8)	18,256	184,496	389	239	150
Physical activity, MET-h/d	28.7(7.3)	27,284	300,982	523	314	209	26.8 (7.0)	17,112	173,159	361	219	142
Alcohol consumption												
Never	23.5	6,355	68,967	96	60	36	23.2	4,192	41,652	83	51	32
Occasional	7.7	2,087	23,652	26	15	11	8.6	1,565	16,013	22	10	12
Regular: <300 g/w	48.1	13,038	143,999	248	155	93	35.4	6,403	65,130	108	64	44
Regular: ≧300 g/w	20.8	5,623	62,184	146	85	61	32.9	5,948	60,187	171	111	60
Smoking status												
Never	23.6	6,579	74,342	111	64	47	36.1	6,483	66,178	110	68	42
Former	23.9	6,657	73,238	142	89	53	16.2	2,901	29,256	78	57	21
Current	52.5	14,601	159,481	284	174	110	47.7	8,555	85,836	195	112	83

[a] CRC, colorectal cancer; MET, metabolic equivalent.
[b] Cohort II (follow-up: 1993–2005) was used to develop the prediction model.
[c] Cohort I (follow-up: 1995–2005) was to evaluate the prediction model's performance.

risk [36]. The Hosmer–Lemeshow χ^2 test was applied to analyze the difference between the observed and estimated risk by groups of deciles [37]. The ratio of observed and expected CRC events (the sum of individual predicted risk probability in a certain risk category) was used to test the model predictive capability for each risk factor in Cohort I. The 95% CIs for O/E ratio was calculated as $(O/E) \times \exp[\pm 1.96\sqrt{(1/O)}]$; the prediction model underestimated the CRC risk if the O/E ratio was >1, while it overestimated the risk if the O/E ratio was <1 [36].

2.4.3. Simple point score model

A simple point score model (risk sheet) for CRC was developed based on the original prediction model, with the transference of continuous variables of age, BMI, and physical activity into category variables [38,39]. The β coefficients were newly fitted by the Cox model with each of category variables. The first step was to round regression coefficients to scores, and in this analysis, we multiplied coefficients by three, and round them [38,40]. Further, the risk score of each participant was assigned by summing the points from each risk factor present. The score sheets provide comparison 10-year absolute risks for persons of the same age from average and low-risk CRC.

All analyses were conducted using SAS version 9.01 (SAS Inc., Cary, NC, USA).

3. Results

As of December 2005, newly diagnosed cases of CRC were 543 in Cohort II and 389 in Cohort I. In total, 310,059 and 184,496 person-years were observed in the average follow-up periods of 11.0 and of 10.1 years in Cohorts II and I, respectively.

Comparisons of model constructions among the full predictive model and the models with reduced variables were shown in Table 1, in which the reduced multivariate model with age, BMI, physical activity, smoking habit and alcohol consumption was the optimal one (the global test for model non-proportionality, $P = 0.984$, 0.597, and 0.093 for CRC, colon, and rectal cancer, respectively). Numbers of participants, person-years of follow-up, and CRC events, as well as the risk factors of CRC are listed in Table 2. The respective β coefficients and HRs for CRC risk factors obtained from Cox regression of Cohorts II and I, with baseline survival rate at 10-years, are shown in Table 3. Risk factors showed similar relationships to CRC, colon, and rectal cancer.

In the discriminatory analysis of Cohort II, the C statistics were 0.70 (95% CI, 0.68–0.72) for CRC, 0.71 (95% CI, 0.68–0.74) for colon cancer, and 0.68 (95% CI, 0.64–0.71) for rectal cancer, showing a good ability to distinguish cases from non-cases. In Cohort I, the C statistics were 0.64 (95% CI, 0.61–0.67) for CRC, 0.66 (95% CI: 0.62–0.70) for colon cancer, and 0.62 (95% CI: 0.57–0.66) for rectal cancer, showing a modest ability to distinguish cases from non-cases.

In the calibration analysis, χ^2 was 14.2 ($P = 0.08$) for CRC, 11.0 ($P = 0.20$) for colon, and 11.2 ($P = 0.19$) for rectum cancer, showing that the actual rates of CRC in Cohort I were similar to the rates predicted by the Cohort II function (Fig. 1). The overall O/E ratios were 1.09 (95% CI, 0.98–1.23) for CRC, 1.19 (95% CI, 1.03–1.37) for colon cancer, and 0.94 (95% CI, 0.78–1.12) for rectal cancer. Agreement between the predicted and the observed number of events was good in most risk factor categories with several exceptions (e.g., underestimation for CRC in the "never" alcohol consumption category and overestimation for rectal cancer in the age group of 45–49) (Table 4).

In addition, when participants who had a history of diabetes (1991 in Cohort II and 1332 in Cohort I) or a family history of CRC in first-degree relatives (475 in Cohort II and 157 in Cohort I) were excluded, the same predictive risk factors were identified, and similar discrimination and calibration values were observed for CRC, colon, and rectal cancer, respectively, in Cohort I (data not shown).

The simple point score model (risk sheet) was developed for CRC in Cohort II (Fig. 2), for which the C statistic was 0.69 (95% CI, 0.67–0.71). In Fig. 2, the average and the lowest risk probability by age groups in Cohort II are also shown. Correspondingly, validation was performed in Cohort I for the simple point score model: the C statistic was 0.61 (95% CI, 0.58–0.64) for CRC, with similar O/E ratios and 95% CIs in each category of risk factors (data not shown).

4. Discussion

We developed a CRC risk prediction model with established risk factors of age, BMI, alcohol consumption, smoking status, and physical activity level for middle-aged Japanese men. The prediction model was well calibrated in an external cohort. We also presented a simple point score model (risk sheet) for CRC risk estimation.

Cancer is a multifactorial disease involving a variety of factors in the development of clinical manifestations. This recognition has

Table 3
β-Coefficients and hazard ratios with 95% confidence intervals of colorectal cancer risk factors in men, Japan Public Health Center-based Prospective Study, 1993–2005[a].

Risk factor	Cohort II[b,d]						Cohort I[c,d]					
	CRC		Colon		Rectum		CRC		Colon		Rectum	
	β	HR (95% CI)	β	HR (95% CI)	β	HR (95%CI)	β	HR (95% CI)	β	HR (95% CI)	β	HR (95% CI)
Age, year	0.080	1.08 (1.07–1.10)	0.085	1.09 (1.07–1.11)	0.067	1.07 (1.05–1.09)	0.063	1.07 (1.05–1.09)	0.062	1.06 (1.04–1.09)	0.065	1.07 (1.04–1.10)
BMI, kg/m^2	0.047	1.05 (1.02–1.08)	0.049	1.05 (1.01–1.09)	–	–	0.003	1.00 (0.97–1.04)	0.013	1.01 (0.97–1.06)	–	–
Physical activity, MET-h/d	−0.019	0.98 (0.97–0.99)	−0.020	0.98 (0.97–1.00)	−0.020	0.98 (0.97–1.00)	−0.017	0.98 (0.97–1.00)	−0.027	0.97 (0.95–0.99)	−0.006	0.99 (0.97–1.02)
Alcohol consumption												
Never	−0.163	0.85 (0.56–1.28)	−0.140	0.87 (0.51–1.49)	−0.149	0.86 (0.48–1.55)	0.314	1.37 (0.88–2.14)	0.474	1.61 (0.88–2.92)	−0.028	0.97 (0.51–1.87)
Occasional		1.00		1.00		1.00		1.00		1.00		1.00
Regular: <300 g/w	0.358	1.43 (0.98–2.09)	0.419	1.52 (0.93–2.50)	0.309	1.36 (0.80–2.31)	0.072	1.07 (0.69–1.67)	0.182	1.20 (0.67–2.16)	−0.197	0.82 (0.43–1.55)
Regular: ≧300 g/w	0.659	1.93 (1.32–2.83)	0.655	1.93 (1.16–3.19)	0.745	2.11 (1.21–3.65)	0.679	1.97 (1.30–3.00)	0.858	2.36 (1.35–4.14)	0.348	1.42 (0.76–2.63)
Smoking status												
Never		1.00		1.00				1.00		1.00		
Former	0.071	1.07 (0.83–1.39)	0.186	1.21 (0.86–1.69)	–	–	0.438	1.55 (1.15–2.09)	0.605	1.83 (1.27–2.64)	–	–
Current	0.239	1.27 (1.01–1.60)	0.347	1.41 (1.04–1.92)	–	–	0.323	1.38 (1.08–1.77)	0.222	1.25 (0.91–1.72)	–	–
Baseline survival function at 10-year, St(10)	0.9882		0.9928		0.9954		0.9835		0.9890		0.9942	

[a] CRC, colorectal cancer; HR, hazard ratio; CI, confidential interval; MET, metabolic equivalent.
[b] Cohort II (follow-up: 1993–2005) was used to develop the prediction model.
[c] Cohort I (follow-up: 1995–2005) was to evaluate the prediction model's performance.
[d] The HR of each risk factor/category was not significantly different between Cohort II and Cohort I ($P > 0.05$) for the model of CRC, colon, and rectal cancer, respectively.

Fig. 1. The 10-year observed and predicted colorectal cancer events in Cohort I men, Japan Public Health Center-based Prospective Study, 1993–2005.

led the development of risk assessment tools that attempt to synthesize the values of numerous variables into a single statement about the risk of developing a cancer [41]. In this prediction model, age, alcohol consumption, and daily physical activity level were identified as the most important CRC risk factors, consistent with other reports [4,18,20]. Although body weight was also a potential predictor in this analysis, BMI was arbitrarily selected in the model building as a relevant comprehensive risk factor of CRC [10,18,20].

Dietary factors such as consumption of red meat, green vegetables, fibers, dairy, calcium supplement use, or intake of folate were not identified in this population, although they were previously reported as possibly related to CRC risk [4,18,42]. Moreover, no dietary food combinations, including total meat (pork, beef, bacon, ham, and sausage) [42], processed meat (bacon, ham, and sausage) [42,43], total white meat (fish and poultry) [42], ratio of red meat to vegetable, or ratio of red meat to white meat [44] were risk predictors of CRC in this study population. Although in recent years the dietary pattern in the Japanese population has tended toward the western pattern, the traditional dietary habits were substantially maintained, especially in older people [45]. This may account for the lack of foods or dietary nutrients serving as significant factors for predicting CRC in men. Alternatively, it might be possible that data in this study population were insufficient to support a quantitative statement about the exact magnitude of risk from these diets.

A previous CRC risk prediction model was developed by means of larger case–control studies and included CRC screening during the previous 3 years and number of relatives with CRC [18]. In our study, sigmoidoscopy/colonoscopy and fecal occult blood test were not available in the Cohort II questionnaire, although these are known as indicators for the secondary prevention for CRC [46]. The personal history of diabetes was reported as a possible risk factor of CRC [26]. In the present study, however, diabetes showed statistical significance for colon cancer in the univariate analysis but not in the multivariate analysis. In addition, few participants reported a family history of CRC, such that this factor could not be considered for entering into the prediction model. In the analysis for participants without history of diabetes or family history of CRC, a similar predictive ability for CRC was observed. This may indicate that these two factors were not powerful enough for prediction of CRC in this population. Nevertheless, most CRC risk factors included in this prediction model represent lifestyle choices that can be modified with the aim of preventing the disease.

Several validation studies on cancer risk prediction models also showed modest discriminatory accuracy as measured by C

Table 4
10-Years of observed and expected colorectal cancer events, ratios and 95% confidential intervals in Cohort I men, Japan Public Health Center-based Prospective Study, 1993–2005[a].

	CRC					Colon					Rectum				
	Observed	Expected	*O/E* ratio	95%	CI	Observed	Expected	*O/E* ratio	95	%CI	Observed	Expected	*O/E* ratio	95%	CI
Overall	322	294	1.09	0.98	1.23	215	181	1.19	1.03	1.37	107	114	0.94	0.78	1.12
Age, years															
45–49	45	39.0	1.15	0.84	1.58	35	22.8	1.53	1.02	2.31	10	16.4	0.61	0.38	0.99
50–54	62	53.2	1.17	0.89	1.53	41	31.8	1.29	0.91	1.82	21	21.4	0.98	0.64	1.50
55–59	95	76.1	1.25	1.00	1.56	55	46.7	1.18	0.88	1.57	40	29.5	1.36	0.95	1.95
60–64	112	119.9	0.93	0.78	1.12	78	75.9	1.03	0.82	1.29	34	44.7	0.76	0.57	1.02
65–69	8	6.2	1.30	0.59	2.86	6	4.0	1.52	0.57	4.07	2	2.3	0.87	0.24	3.14
BMI, kg/m^2															
<25	230	200.9	1.14	1.00	1.31	153	123.6	1.24	1.04	1.48	–	–	–	–	–
≧25	92	93.5	0.98	0.80	1.21	62	57.6	1.08	0.83	1.39	–	–	–	–	–
Physical activity, MET-h/d															
<22.0	118	109.3	1.08	0.89	1.30	92	67.8	1.36	1.07	1.72	33	41.9	0.79	0.58	1.07
22.0-<28.9	95	101.4	0.94	0.77	1.14	70	62.4	1.12	0.87	1.44	34	39.4	0.86	0.63	1.18
≧28.9	83	83.6	0.99	0.80	1.23	57	50.9	1.12	0.85	1.47	33	33.1	1.00	0.71	1.40
Alcohol consumption															
Never	66	42.5	1.55	1.15	2.10	48	26.0	1.84	1.26	2.71	18	16.5	1.09	0.67	1.77
Occasional	19	17.7	1.07	0.67	1.71	9	10.6	0.85	0.47	1.56	10	6.4	1.57	0.72	3.42
Regular: <300 g/w	95	103.0	0.92	0.76	1.12	59	65.5	0.90	0.71	1.15	36	37.9	0.95	0.69	1.31
Regular: ≧300 g/w	137	129.6	1.06	0.89	1.26	96	78.2	1.23	0.98	1.53	41	53.1	0.77	0.59	1.01
Smoking status															
Never	87	91.6	0.95	0.77	1.17	58	52.7	1.10	0.84	1.44	–	–	–	–	–
Former	69	48.9	1.41	1.07	1.87	52	31.5	1.65	1.16	2.34	–	–	–	–	–
Current	160	149.7	1.07	0.91	1.25	103	94.6	1.09	0.89	1.33	–	–	–	–	–

[a] CRC, colorectal cancer; *O/E*, observed/expected; CI, confidential interval; MET, metabolic equivalent.

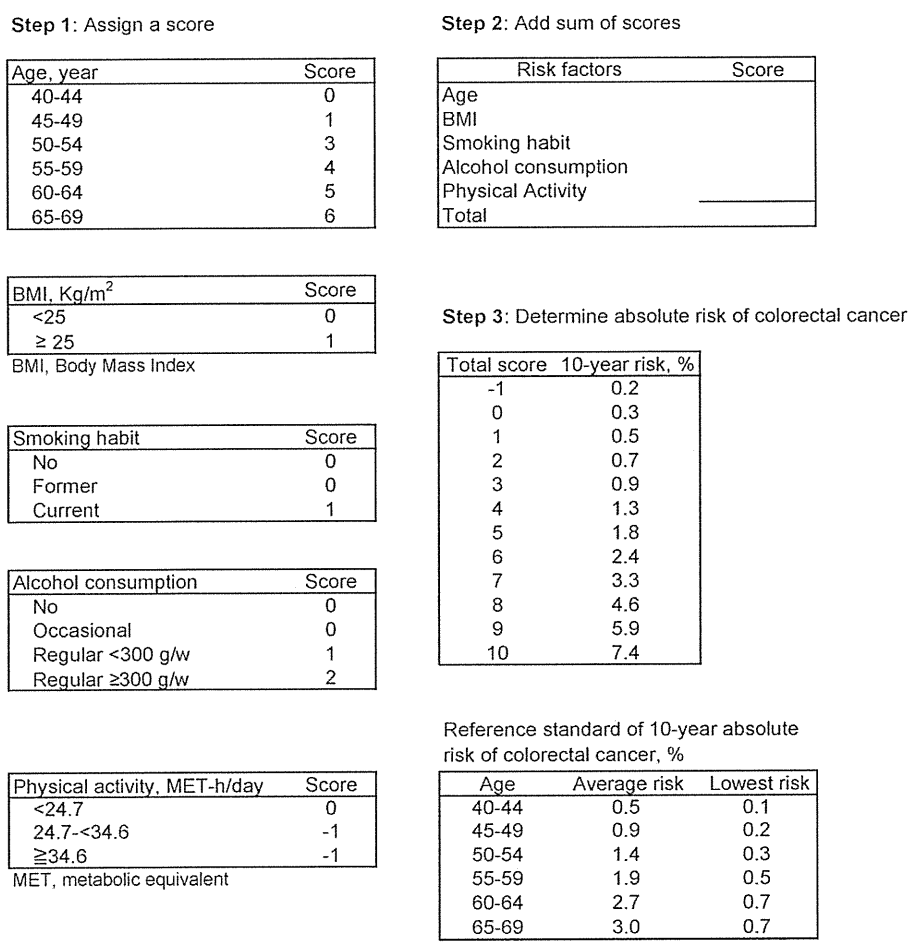

Fig. 2. Simple point score model (risk sheet) for evaluation of 10-year risk of colorectal cancer incidence in men.

statistics, including 0.61 for CRC [36], 0.60–0.63 for breast cancer [47,48], and 0.60–0.69 for lung cancer [49,50]. Similarly, the modest ability to predict CRC in this study suggested that in future studies stronger risk predictors need to be found [18], for instance, dietary nutrient intake or genotypes.

The overall predicted number of CRC events was close to the actual number, with several exceptions in the validation. The differences between the observed and the predicted CRC events in Cohort I may be due to a different distribution of participants with higher risk in the two cohorts. For example, more elderly men and smokers were in Cohort II than in Cohort I, while more heavy alcohol drinkers were in Cohort I than in Cohort II. The discrepancies in the questionnaires used in the two cohorts also may partly account for the difference [36].

The validation in this study was done in an external cohort (Cohort I); however, risk factor profiles and measurement were similar to those of the population for model development (Cohort II). Therefore, the generalizability of the prediction model needs to be tested in other populations to provide more external validations. Another limitation of this study was that the simple point score model (risk sheet) for estimation of CRC risk included not only simple frequency components (age, body weight, and smoking) but also those based on calculation (alcohol consumption by gram per week and physical activity by MET-hour per day). This may make it inconvenient for an individual to use the sheet directly. In addition, because the 5-year follow-up measurement was used as the baseline for Cohort I in this analysis, the smaller relevant population might reduce its validation capability.

In summary, the CRC risk prediction model was developed based on a large cohort study; it showed modest discrimination power and was well calibrated in another large cohort. This model may be used by clinicians, public health professionals, and individuals to estimate the CRC risk for Japanese men, which could play a role in the promotion of CRC prevention strategies. Further validation in other populations, with the addition of more established factors, is necessary.

Conflict of interest statement

None declared.

Acknowledgments

This study was supported by a Grant-in-Aid for Cancer Research (19 shi-2) and for the Third Term Comprehensive 10-year Strategy for Cancer Control (H21-Sanjigan-Ippan-003) from the Ministry of Health, Labour and Welfare of Japan and Grants-in-Aid for Scientific Research for Young Scientists (A) (19689014) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and Japan Society for the Promotion of Science.

We thank all staff members in each study area for their painstaking efforts to conduct the baseline survey and follow-up. Members of the JPHC Study Group (principal investigator: S. Tsugane) include: S. Tsugane, M. Inoue, T. Sobue, and T. Hanaoka, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo; J. Ogata, S. Baba, T. Mannami, A. Okayama, and Y. Kokubo, National Cardiovascular Center, Suita; K. Miyakawa, F. Saito, A. Koizumi, Y. Sano, I. Hashimoto, T. Ikuta, and Y. Tanaba, Iwate Prefectural Ninohe Public Health Center, Ninohe; Y. Miyajima, N. Suzuki, S. Nagasawa, Y. Furusugi, and N. Nagai, Akita Prefectural Yokote Public Health Center, Yokote; H. Sanada, Y. Hatayama, F. Kobayashi, H. Uchino, Y. Shirai, T. Kondo, R. Sasaki, Y. Watanabe, Y. Miyagawa, Y. Kobayashi, and M. Machida, Nagano Prefectural Saku Public Health Center, Saku; Y. Kishimoto, E. Takara, T. Fukuyama, M. Kinjo, M. Irei, and H. Sakiyama, Okinawa Prefectural Chubu Public Health Center, Okinawa; K. Imoto, H. Yazawa, T. Seo, A. Seiko, F. Ito, F. Shoji, and R. Saito, Katsushika Public Health Center, Tokyo; A. Murata, K. Minato, K. Motegi, and T. Fujieda, Ibaraki Prefectural Mito Public Health Center, Mito; T. Abe, M. Katagiri, M. Suzuki, and K. Matsui, Niigata Prefectural Kashiwazaki and Nagaoka Public Health Center, Kashiwazaki and Nagaoka; M. Doi, A. Terao, Y. Ishikawa, and T. Tagami, Kochi Prefectural Chuo-higashi Public Health Center, Tosayamada; H. Doi, M. Urata, N. Okamoto, F. Ide, and H. Sueta, Nagasaki Prefectural Kamigoto Public Health Center, Arikawa; H. Sakiyama, N. Onga, H. Takaesu, and M. Uehara, Okinawa Prefectural Miyako Public Health Center, Hirara; F. Horii, I. Asano, H. Yamaguchi, K. Aoki, S. Maruyama, M. Ichii, and M. Takano, Osaka Prefectural Suita Public Health Center, Suita; S. Matsushima and S. Natsukawa, Saku General Hospital, Usuda; M. Akabane, Tokyo University of Agriculture, Tokyo; M. Konishi, K. Okada, and I. Saito, Ehime University, Toon; H. Iso, Osaka University, Suita; Y. Honda, K. Yamagishi, S. Sakurai, and N. Tsuchiya, Tsukuba University, Tsukuba; H. Sugimura, Hamamatsu University, Hamamatsu; Y. Tsubono, Tohoku University, Sendai; M. Kabuto, National Institute for Environmental Studies, Tsukuba; S. Tominaga, Aichi Cancer Center Research Institute, Nagoya; M. Iida, W. Ajiki, and A. Ioka, Osaka Medical Center for Cancer and Cardiovascular Disease, Osaka; S. Sato, Osaka Medical Center for Health Science and Promotion, Osaka; N. Yasuda, Kochi University, Nankoku; K. Nakamura, Niigata University, Niigata; S. Kono, Kyushu University, Fukuoka; K. Suzuki, Research Institute for Brain and Blood Vessels Akita, Akita; Y. Takashima and M. Yoshida, Kyorin University, Mitaka; E. Maruyama, Kobe University, Kobe; M. Yamaguchi, Y. Matsumura, S. Sasaki, and S. Watanabe, National Institute of Health and Nutrition, Tokyo; T. Kadowaki, Tokyo University, Tokyo; M. Noda and T. Mizoue, International Medical Center of Japan, Tokyo; Y. Kawaguchi, Tokyo Medical and Dental University, Tokyo; H. Shimizu, Sakihae Institute, Gifu.

References

[1] Matsuda T, Marugame T, Kamo K, Katanoda K, Ajiki W, Sobue T. Cancer incidence and incidence rates in Japan in 2002: based on data from 11 population-based cancer registries. Jpn J Clin Oncol 2008;38:641–8.
[2] The Editorial Board of the Cancer Statistics in Japan, ed. Cancer statistics in Japan 2007. Tokyo: Foundation for Promotion of Cancer Research (FPCR), 2007
[3] Shibuya K, Mathers CD, Boschi-Pinto C, Lopez AD, Murray CJ. Global and regional estimates of cancer mortality and incidence by site. II. Results for the global burden of disease 2000. BMC Cancer 2002;2:37.
[4] Wiseman M. The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Proc Nutr Soc 2008;67:253–6.
[5] Botteri E, Iodice S, Bagnardi V, Raimondi S, Lowenfels AB, Maisonneuve P. Smoking and colorectal cancer: a meta-analysis. JAMA 2008;300:2765–78.
[6] Lee KJ, Inoue M, Otani T, Iwasaki M, Sasazuki S, Tsugane S. Physical activity and risk of colorectal cancer in Japanese men and women: the Japan Public Health Center-based prospective study. Cancer Causes Contr 2007;18:199–209.
[7] Isomura K, Kono S, Moore MA, Toyomura K, Nagano J, Mizoue T, et al. Physical activity and colorectal cancer: the Fukuoka colorectal cancer study. Cancer Sci 2006;97:1099–104.
[8] Otani T, Iwasaki M, Yamamoto S, Sobue T, Hanaoka T, Inoue M, et al. Alcohol consumption, smoking, and subsequent risk of colorectal cancer in middle-aged and elderly Japanese men and women: Japan Public Health Center-based prospective study. Cancer Epidemiol Biomarkers Prev 2003;12:1492–500.
[9] Shimizu N, Nagata C, Shimizu H, Kametani M, Takeyama N, Ohnuma T, et al. Height, weight, and alcohol consumption in relation to the risk of colorectal cancer in Japan: a prospective study. Br J Cancer 2003;88:1038–43.
[10] Otani T, Iwasaki M, Inoue M. Body mass index, body height, and subsequent risk of colorectal cancer in middle-aged and elderly Japanese men and women: Japan public health center-based prospective study. Cancer Causes Contr 2005;16:839–50.
[11] Tsubono Y, Otani T, Kobayashi M, Yamamoto S, Sobue T, Tsugane S. No association between fruit or vegetable consumption and the risk of colorectal cancer in Japan. Br J Cancer 2005;92:1782–4.
[12] Otani T, Iwasaki M, Ishihara J, Sasazuki S, Inoue M, Tsugane S. Dietary fiber intake and subsequent risk of colorectal cancer: the Japan Public Health Center-based prospective study. Int J Cancer 2006;119:1475–80.
[13] Mizoue T, Inoue M, Wakai K, Nagata C, Shimazu T, Tsuji I, et al. Alcohol drinking and colorectal cancer in Japanese: a pooled analysis of results from five cohort studies. Am J Epidemiol 2008;167:1397–406.

[14] Mizoue T, Inoue M, Tanaka K, Tsuji I, Wakai K, Nagata C, et al. Tobacco smoking and colorectal cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population. Jpn J Clin Oncol 2006;36:25–39.
[15] Akhter M, Inoue M, Kurahashi N, Iwasaki M, Sasazuki S, Tsugane S. Reproductive factors, exogenous female hormone use and colorectal cancer risk: the Japan Public Health Center-based Prospective Study. Eur J Cancer Prev 2008;17:515–24.
[16] Tamakoshi K, Wakai K, Kojima M, Watanabe Y, Hayakawa N, Toyoshima H, et al. A prospective study of reproductive and menstrual factors and colon cancer risk in Japanese women: findings from the JACC study. Cancer Sci 2004;95:602–7.
[17] Tsugane S. What we know about associations between diet and cancer. JMAJ 2008;51:7.
[18] Freedman AN, Slattery ML, Ballard-Barbash R, Willis G, Cann BJ, Pee D, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. J Clin Oncol 2008.
[19] Parkin DM, Olsen AH, Sasieni P. The potential for prevention of colorectal cancer in the UK. Eur J Cancer Prev 2009;18:179–90.
[20] Colditz GA, Atwood KA, Emmons K, Monson RR, Willett WC, Trichopoulos D, et al. Harvard report on cancer prevention volume 4: Harvard cancer risk index. Risk Index Working Group, Harvard Center for Cancer Prevention. Cancer Causes Contr 2000;11:477–88.
[21] Selvachandran SN, Hodder RJ, Ballal MS, Jones P, Cade D. Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: a prospective study. Lancet 2002;360:278–83.
[22] Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. Ann Intern Med 2003;139:959–65.
[23] Tsugane S, Sobue T. Baseline survey of JPHC study—design and participation rate. Japan public health center-based prospective study on cancer and cardiovascular diseases. J Epidemiol 2001;11:S24–9.
[24] Iwasaki M, Otani T, Yamamoto S, Inoue M, Hanaoka T, Sobue T, et al. Background characteristics of basic health examination participants: the JPHC study baseline survey. J Epidemiol 2003;13:216–25.
[25] Sasaki S, Kobayashi M, Ishihara J, Tsugane S. Self-administered food frequency questionnaire used in the 5-year follow-up survey of the JPHC study: questionnaire structure, computation algorithms, and area-based mean intake. J Epidemiol/Jpn Epidemiol Assoc 2003;13:S13–22.
[26] Inoue M, Iwasaki M, Otani T, Sasazuki S, Noda M, Tsugane S. Diabetes mellitus and the risk of cancer: results from a large-scale population-based cohort study in Japan. Arch Intern Med 2006;166:1871–7.
[27] Inoue M, Iso H, Yamamoto S, Kurahashi N, Iwasaki M, Sasazuki S, et al. Daily total physical activity level and premature death in men and women: results from a large-scale population-based cohort study in Japan (JPHC study). Ann Epidemiol 2008;18:522–30.
[28] Technology. CfSa, Ministry of Education C., Sports, Science and Technology the Government of Japan, ed. Standard tables of food composition in Japan, the fifth revised and enlarged edition. Tokyo: Printing Bureau, Ministry of Finance, 2005.
[29] Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.
[30] Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 1999;28:964–74.
[31] Sauerbrei W, Meier-Hirmer C, Benner A, Royston P. Multivariable regression model building by using fractional polynomials: description of SAS, STATA and R programs. Comput Stat Data Anal 2006;50:3464–85.
[32] Liu J, Hong Y, D'Agostino Sr RB, Wu Z, Wang W, Sun J, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. JAMA 2004;291:2591–9.
[33] D'Agostino Sr RB, Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001;286:180–7.
[34] Therneau TM, Grambsch GP. Expected survival. Modeling survival data: extending the Cox model. Springer; 2004. p. 280.
[35] D'Agostino RB, Nam BH. Evaluation of the performance of survival analysis models: discrimination and calibration measures. In: Balakrishnan NRC, ed. Handbook of statistics, vol. 23. London, England: Elsevier, 2004.
[36] Park Y, Freedman AN, Gail MH, Pee D, Hollenbeck A, Schatzkin A, et al. Validation of a colorectal cancer risk prediction model among white patients age 50 years and older. J Clin Oncol 2009;27:694–8.
[37] Lemeshow S, Hosmer Jr DW. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol 1982;115:92–106.
[38] Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837–47.
[39] Wu Y, Liu X, Li X, Li Y, Zhao L, Chen Z, et al. Estimation of 10-year risk of fatal and nonfatal ischemic cardiovascular diseases in Chinese adults. Circulation 2006;114:2217–25.
[40] Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. Springer; 2009.
[41] Kannel WB, McGee DL. Composite scoring—methods and predictive validity: insights from the Framingham Study. Health Serv Res 1987;22:499–535.
[42] Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC. Intake of fat, meat, and fiber in relation to risk of colon cancer in men. Cancer Res 1994;54:2390–7.
[43] Cross AJ, Leitzmann MF, Gail MH, Hollenbeck AR, Schatzkin A, Sinha R. A prospective study of red and processed meat intake in relation to cancer risk. PLoS Med 2007;4:e325.
[44] McCullough ML, Feskanich D, Stampfer MJ, Giovannucci EL, Rimm EB, Hu FB, et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. Am J Clin Nutr 2002;76:1261–71.
[45] Kim MK, Sasaki S, Sasazuki S, Tsugane S. Prospective study of three major dietary patterns and risk of gastric cancer in Japan. Int J Cancer 2004;110:435–42.
[46] Lee KJ, Inoue M, Otani T, Iwasaki M, Sasazuki S, Tsugane S. Colorectal cancer screening using fecal occult blood test and subsequent risk of colorectal cancer: a prospective cohort study in Japan. Cancer Detect Prev 2007;31:3–11.
[47] Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J Natl Cancer Inst 2006;98:1215–26.
[48] Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst 2006;98:1204–14.
[49] Rosner BA, Colditz GA, Webb PM, Hankinson SE. Mathematical models of ovarian cancer incidence. Epidemiology 2005;16:508–15.
[50] Cronin KA, Gail MH, Zou Z, Bach PB, Virtamo J, Albanes D. Validation of a model of lung cancer risk prediction among smokers. J Natl Cancer Inst 2006;98:637–40.