Table 1. Comparison of daily intakes of energy and 26 nutrients measured with three-day weighed diet records vs. food frequency questionnaire.

		Male (n=73)						Female (n=129)					
		3d-WDRs		FFQ		Ratio of FFQ to		3d-WDRs		FFQ		Ratio of FFQ to	
Nutrient		Mean	SD	Mean	SD	3d-WDRs	p	Mean	SD	Mean	SD	3d-WDRs	p
Energy	[kcal]	2342	469	1987	268	0.85	***	1924	332	1639	186	0.85	***
Protein	[g]	88.4	22.1	60.8	10.2	0.69	***	74.5	16.3	55.2	7.8	0.74	***
Fat	[g]	66.1	22.6	47.1	11.9	0.71	***	59.2	16.5	48.4	9.6	0.82	***
Carbohydrate	[g]	312.7	57.7	293.0	51.7	0.94	*	264.5	50.0	226.6	36.1	0.86	***
Protein energy% †	[%]	15.1	2.0	12.3	1.4	0.81	***	15.5	2.0	13.5	1.5	0.87	***
Fat energy% †	[%]	25.1	5.4	21.4	4.6	0.85	***	27.5	5.1	26.7	4.9	0.97	
Carbohydrate energy% †	[%]	53.9	6.2	58.8	4.6	1.09	***	55.2	6.1	55.2	5.0	1.00	
Saturated fatty acids	[g]	16.6	6.6	11.3	2.0	0.68	***	16.0	5.5	12.4	2.5	0.78	***
Monounsaturated fatty acids	[g]	23.1	9.3	17.5	4.4	0.76	***	19.8	6.2	16.9	3.4	0.85	***
Polyunsaturated fatty acids	[g]	16.4	5.3	14.1	3.2	0.86	**	14.0	4.1	13.5	2.9	0.97	
n-6 Polyunsaturated fatty acids	[g]	12.8	4.5	11.8	2.7	0.92	**	11.0	3.4	11.5	2.6	1.04	
n-3 Polyunsaturated fatty acids	[g]	3.3	1.2	2.3	0.5	0.70	***	2.8	1.1	2.2	0.5	0.80	***
n-3 Highly-unsaturated fatty acids	[g]	1.1	0.9	0.7	0.3	0.66	***	0.9	0.7	0.7	0.2	0.78	
Cholesterol	[mg]	424	176	274	64	0.65	***	345	132	264	64	0.76	***
Iron	[mg]	9.8	2.4	7.7	1.9	0.79	***	8.9	2.7	7.7	1.6	0.86	***
Calcium	[mg]	592	186	508	129	0.86	*	609	231	566	144	0.93	
Carotene	[μ g]	4244	1840	3229	1285	0.76	*	4241	2103	3550	1131	0.84	***
Vitamin A	[μ gRE]	989	478	1052	384	1.06		1067	832	1052	422	0.99	
Vitamin D	[μ g]	9.4	5.4	7.4	3.4	0.79	**	8.0	5.9	7.2	2.6	0.91	
Vitamin E	[mg α -TE]	10.1	3.3	8.6	2.1	0.85	**	9.4	3.0	8.6	1.8	0.92	*
Vitamin B_1	[mg]	1.18	0.4	0.69	0.08	0.58	***	1.04	0.30	0.70	0.10	0.65	***
Vitamin B_2	[mg]	1.48	0.44	1.12	0.21	0.76	**	1.38	0.43	1.20	0.20	0.89	***
Folate	[μ g]	417	148	357	109	0.86	**	409	164	384	93	0.94	
Vitamin C	[mg]	123	57	103	34	0.84		136	69	122	34	0.90	
Soluble dietary fiber	[g]	3.7	1.2	2.1	0.6	0.57	***	2.4	0.7	2.3	0.5	0.61	***
Insoluble dietary fiber	[g]	12.1	3.2	8.0	2.2	0.66	***	12.0	3.7	9.0	1.9	0.75	***
Total dietary fiber	[g]	16.6	4.4	11.4	3.1	0.69	***	16.6	5.1	12.4	2.7	0.75	***
Median						0.79						0.86	
Average						0.79						0.85	

*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$.

†: Percentage of energy from protein, fat or carbohydrate to total energy.

3d-WDRs: 3-day weighed diet records, FFQ: food frequency questionnaire, SD: standard deviation.

Table 2. Pearson's and Spearman's rank correlation coefficients (CCs) between intakes of energy and 26 nutrients measured with three-day weighed diet records and food frequency questionnaire for males.

	Pearson's CCs*					Spearman's rank CCs*	
Nutrient	Crude	Log-transformed †	Log- transformed and energy- adjusted ‡	σ^2w/σ^2b §	De-attenuated, log-transformed and energy-adjusted | (95% CI)	Crude	Energy-adjusted
Energy	0.41	0.40		1.4	0.49 (0.29 - 0.65) ¶	0.36	
Protein	0.36	0.32	0.42	1.3	0.50 (0.25 - 0.71)	0.22	0.35
Fat	0.53	0.48	0.52	1.2	0.62 (0.39 - 0.80)	0.38	0.49
Carbohydrate	0.54	0.55	0.73	1.1	0.86 (0.71 - 0.96)	0.57	0.73
Protein energy%	0.45	0.45	0.42	1.3	0.51 (0.26 - 0.72)	0.38	0.35
Fat energy%	0.55	0.56	0.51	1.2	0.61 (0.38 - 0.79)	0.49	0.50
Carbohydrate energy%	0.68	0.70	0.74	1.1	0.86 (0.71 - 0.97)	0.68	0.76
Saturated fatty acids	0.50	0.43	0.55	1.0	0.64 (0.48 - 0.90)	0.35	0.52
Monounsaturated fatty acids	0.52	0.44	0.37	1.2	0.43 (0.15 - 0.55)	0.12	0.32
Polyunsaturated fatty acids	0.35	0.31	0.34	1.9	0.44 (0.14 - 0.61)	0.05	0.33
n-6 Polyunsaturated fatty acids	0.20	0.21	0.10	1.6	0.12 (-0.11 - 0.34)	0.20	0.13
n-3 Polyunsaturated fatty acids	0.35	0.36	0.41	2.5	0.55 (0.37 - 0.69)	0.37	0.37
n-3 Highly-unsaturated fatty acids	0.14	0.31	0.28	2.1	0.36 (0.14 - 0.55)	0.28	0.23
Cholesterol	0.35	0.25	0.10	2.1	0.13 (-0.16 - 0.38)	0.15	0.15
Calcium	0.32	0.34	0.42	1.0	0.49 (0.25 - 0.69)	0.38	0.43
Iron	0.25	0.26	0.49	1.2	0.58 (0.35 - 0.76)	0.21	0.50
Carotene	0.19	0.23	0.29	2.18	0.39 (0.09 - 0.65)	0.18	0.28
Vitamin A	0.18	0.16	0.21	2.25	0.27 (-0.03 - 0.55)	0.10	0.19
Vitamin D	0.34	0.40	0.45	3.21	0.65 (0.36 - 0.89)	0.33	0.35
Vitamin E	0.25	0.21	0.25	1.83	0.31 (0.02 - 0.57)	0.16	0.27
Vitamin B_1	0.31	0.25	0.21	1.73	0.26 (-0.03 - 0.52)	0.19	0.19
Vitamin B_2	0.35	0.31	0.48	1.11	0.57 (0.36 - 0.77)	0.34	0.53
Folate	0.12	0.17	0.33	0.72	0.36 (0.12 - 0.58)	0.21	0.41
Vitamin C	0.27	0.27	0.40	0.94	0.45 (0.21 - 0.66)	0.24	0.52
Soluble dietary fiber	0.04	0.07	0.21	1.38	0.25 (-0.03 - 0.50)	0.28	0.20
Insoluble dietary fiber	0.11	0.10	0.27	1.53	0.33 (0.06 - 0.58)	0.22	0.24
Total dietary fiber	0.12	0.12	0.30	1.44	0.36 (0.09 - 0.60)	0.34	0.27
Median	0.34	0.31	0.38	1.36	0.45	0.46	0.35
Average	0.33	0.32	0.38	1.54	0.46	0.29	0.37

* : For n=73, $r > 0.24$ ($p<0.05$), $r > 0.31$ ($p<0.01$), $r > 0.39$ ($p<0.001$).

† : All energy and nutrients intakes were $\log_e$- transformed to improve normality.

‡ : Energy intake was adjusted using residual model.

§ : Ratio of within-person to between-person variance of nutrient intakes from three-day weighed diet records.

| : De-attenuated correlation coefficient is calculated using ratio of within- to between-person variation measured with three-day weighed diet records.

¶ : De-attenuation only.

CI: confidence interval.

Table 3. Pearson's and Spearman's rank correlation coefficients between intakes of energy and 26 nutrients measured with three-day weighed diet records and food frequency questionnaire for females.

	Pearson's CCs*					Spearman's rank CCs*	
Nutrient	Crude	Log-transformed †	Log- transformed and energy- adjusted ‡	σ^2w/σ^2b §	De-attenuated, log-transformed and energy-adjusted ‖ (95% CI)	Crude	Energy-adjusted
Energy	0.38	0.38		0.97	0.44 (0.30 - 0.65) ¶	0.37	
Protein	0.31	0.31	0.29	1.60	0.36 (0.25 - 0.62) **	0.30	0.33
Fat	0.29	0.29	0.40	1.32	0.48 (0.40 - 0.72)	0.22	0.38
Carbohydrate	0.48	0.52	0.55	1.05	0.64 (0.61 - 0.85)	0.45	0.44
Protein energy%	0.33	0.33	0.30	1.61	0.37 (0.26 - 0.63)	0.37	0.34
Fat energy%	0.36	0.37	0.40	1.33	0.48 (0.40 - 0.72)	0.33	0.37
Carbohydrate energy%	0.55	0.57	0.57	1.07	0.66 (0.64 - 0.87)	0.45	0.46
Saturated fatty acids	0.40	0.39	0.35	1.33	0.42 (0.32 - 0.66)	0.35	0.34
Monounsaturated fatty acids	0.21	0.18	0.28	1.54	0.34 (0.22 - 0.60)	0.12	0.26
Polyunsaturated fatty acids	0.09	0.13	0.20	1.73	0.25 (0.10 - 0.51)	0.05	0.16
n-6 Polyunsaturated fatty acids	0.11	0.16	0.25	1.60	0.31 (0.14 - 0.46)	0.20	0.22
n-3 Polyunsaturated fatty acids	0.09	0.12	0.17	2.50	0.23 (0.06 - 0.39)	0.17	0.17
n-3 Highly-unsaturated fatty acids	0.17	0.27	0.27	2.10	0.35 (0.19 - 0.49)	0.29	0.27
Cholesterol	0.13	0.15	0.14	2.42	0.19 (0.02 - 0.47)	0.15	0.17
Calcium	0.48	0.52	0.52	0.85	0.59 (0.53 - 0.78)	0.50	0.47
Iron	0.31	0.33	0.38	1.03	0.44 (0.34 - 0.66)	0.33	0.37
Carotene	0.28	0.28	0.30	1.69	0.38 (0.28 - 0.65)	0.31	0.30
Vitamin A	0.11	0.14	0.17	1.90	0.22 (0.06 - 0.48)	0.22	0.24
Vitamin D	0.18	0.27	0.29	2.64	0.40 (0.33 - 0.73)	0.25	0.26
Vitamin E	0.03	0.03	0.14	1.63	0.17 (0.00 - 0.41)	0.00	0.14
Vitamin B_1	0.11	0.09	0.08	2.12	0.10 (-0.10 - 0.35)	0.13	0.11
Vitamin B_2	0.42	0.37	0.37	1.05	0.43 (0.32 - 0.65)	0.38	0.38
Folate	0.25	0.27	0.34	0.84	0.38 (0.25 - 0.59)	0.29	0.36
Vitamin C	0.40	0.40	0.46	0.79	0.52 (0.43 - 0.71)	0.43	0.43
Soluble dietary fiber	0.23	0.26	0.31	1.37	0.37 (0.25 - 0.62)	0.28	0.36
Insoluble dietary fiber	0.31	0.30	0.37	1.35	0.46 (0.36 - 0.70)	0.32	0.37
Total dietary fiber	0.33	0.34	0.40	1.23	0.47 (0.38 - 0.71)	0.34	0.41
Median	0.29	0.29	0.31	1.37	0.38	0.30	0.34
Average	0.27	0.29	0.32	1.51	0.39	0.28	0.31

*: For n=129, $r > 0.20$ ($p<0.05$), $r > 0.26$ ($p<0.01$), $r > 0.32$ ($p<0.001$).
†: All energy and nutrients intakes were $\log_e$- transformed to improve normality.
‡: Energy intake was adjusted using residual model.
§: Ratio of within-person to between-person variance of nutrient intakes from three-day weighed diet records.
‖: De-attenuated correlation coefficient is calculated using ratio of within- to between-person variation measured with three-day weighed diet records.
¶: De-attenuation only.
CI: confidence interval.

Table 4. Comparison of nutrient intakes between three-day weighed diet records and food frequency questionnaire according to quartile classification for males.

	Crude (%)			Energy-adjusted (%)		
Nutrient	Exact agreement	Agreement within adjacent categories	Disagreement	Exact agreement	Agreement within adjacent categories	Disagreement
Energy	33	74	3			
Protein	33	66	8	29	75	5
Fat	32	75	8	42	84	3
Carbohydrate	41	85	0	42	92	0
Protein energy%	29	75	5	32	77	4
Fat energy%	45	79	4	41	79	3
Carbohydrate energy%	51	89	3	49	93	0
Saturated fatty acids	30	75	8	41	85	5
Monounsaturated fatty acids	33	73	7	29	71	4
Polyunsaturated fatty acids	27	71	5	32	74	7
n-6 Polyunsaturated fatty acids	29	71	11	26	62	17
n-3 Polyunsaturated fatty acids	25	74	12	28	71	15
n-3 Highly-unsaturated fatty acids	31	74	6	33	70	9
Cholesterol	32	70	4	25	70	12
Calcium	30	77	5	32	78	3
Iron	30	68	5	42	82	4
Carotene	32	68	10	37	66	10
Vitamin A	29	60	11	27	66	8
Vitamin D	37	74	4	38	75	7
Vitamin E	26	66	11	29	71	7
Vitamin B_1	23	66	7	36	66	5
Vitamin B_2	29	78	3	42	82	1
Folate	30	73	7	38	79	5
Vitamin C	33	67	5	33	74	3
Soluble dietary fiber	23	56	12	32	68	12
Insoluble dietary fiber	22	62	11	33	70	7
Total dietary fiber	26	62	10	26	70	5
Median	30	73	7	33	74	5
Average	31	71	7	34	75	6

Table 5. Comparison of nutrient intakes between three-day weighed diet records and food frequency questionnaire according to quartile classification for females.

	Crude (%)			Energy-adjusted (%)		
Nutrient	Exact agreement	Agreement within adjacent categories	Disagreement	Exact agreement	Agreement within adjacent categories	Disagreement
Energy	31	75	5	33	77	5
Protein	36	73	7	34	75	4
Fat	36	68	9	36	76	6
Carbohydrate	40	76	5	41	78	5
Protein energy%	35	78	5	35	77	3
Fat energy%	33	73	8	37	74	7
Carbohydrate energy%	40	78	5	40	81	5
Saturated fatty acids	33	74	6	39	79	9
Monounsaturated fatty acids	33	68	12	36	72	8
Polyunsaturated fatty acids	26	64	13	27	68	11
n-6 Polyunsaturated fatty acids	29	71	11	26	62	13
n-3 Polyunsaturated fatty acids	25	74	12	28	71	12
n-3 Highly-unsaturated fatty acids	31	74	6	33	70	7
Cholesterol	31	66	11	33	73	12
Calcium	38	81	5	36	83	5
Iron	33	72	7	35	77	5
Carotene	32	77	8	33	73	6
Vitamin A	29	68	6	33	73	9
Vitamin D	32	74	9	29	74	9
Vitamin E	22	63	14	26	67	9
Vitamin B_1	30	67	10	29	65	9
Vitamin B_2	35	76	6	35	75	5
Folate	32	74	9	40	74	7
Vitamin C	39	78	3	36	78	4
Soluble dietary fiber	33	72	5	29	76	4
Insoluble dietary fiber	39	74	9	40	77	5
Total dietary fiber	40	73	7	40	76	5
Median	33	73	7	35	76	6
Average	33	73	8	34	75	7

Table 6. Comparison of validity indices for selected nutrients of Japanese short food frequency questionnaires vs. diet records.

	Takatsuka et al. (1997) *	Egami et al. (1999) †				Tsubono et al. (2001) ‡		Lee et al. (2002) §	Ogawa et al. (2003) ∣		Present study (2004)	
No. of food items	31	97				44		36	21		47	
No. of frequency categories	8	9				4 or 6		5 or 6	5		8	
Procedures of dietary records	24 hour-recall x 12 months	4day WDRs x four seasons				7day-WDRs x 4 seasons		7 day WDRs x 4 seasons	7 consecutive day-WDRs		3 day-WDRs	
Sequence of two methods	24H-Rs →FFQ	FFQ_1 → WDRs		WDRs → FFQ_2		WDRs → FFQ		WDRs → FFQ	WDRs → FFQ		FFQ → WDRs	
Sex	Male and Female	Male	Female	Male	Female	Male	Female	Male	Male	Female	Male	Female
No. of subjects	31	44	42	44	42	94	107	23	55	58	73	129
Energy	0.55	0.25	0.39	0.21	0.38			0.23	0.55	0.38	0.40	0.44
Protein	0.57	0.19	0.30	0.24	0.53	0.39	0.29	0.44	0.25	0.49	0.50	0.36
Fat	-0.03	0.62	0.30	0.60	0.50	0.55	0.37	0.19	0.37	0.50	0.62	0.48
Carbohydrate	0.34	0.52	0.24	0.46	0.53	0.66	0.29	0.45	0.57	0.43	0.86	0.64
Saturated fatty acids	0.51	0.76	0.37	0.73	0.48						0.64	0.42
Monounsaturated fatty acids	0.12	0.61	0.28	0.63	0.53						0.43	0.34
Polyunsaturated fatty acids	-0.15	0.39	0.42	0.39	0.49						0.44	0.25
Cholesterol	0.52	0.53	0.21	0.50	0.35						0.13	0.19
Calcium	0.69	0.61	0.73	0.71	0.78	0.73	0.57	0.52	0.57	0.67	0.49	0.59
Iron		0.22	0.57	0.12	0.52	0.72	0.48	0.31	0.35	0.47	0.58	0.44
Carotene	0.45	0.36	0.51	0.33	0.46	0.52	0.65		0.57	0.45	0.39	0.38
Vitamin A	0.22	0.46	0.52	0.49	0.45			0.19			0.27	0.22
Vitamin D	0.53										0.65	0.40
Vitamin E	0.50	0.50	0.40	0.58	0.41						0.31	0.17
Vitamin B_1						0.44	0.41		0.33	0.31	0.26	0.10
Vitamin B_2						0.60	0.63		0.43	0.54	0.57	0.43
Vitamin C	0.44	0.45	0.40	0.55	0.53	0.49	0.39	0.35	0.58	0.43	0.45	0.52
Soluble dietary fiber								0.52			0.25	0.25
Insoluble dietary fiber								0.56			0.33	0.33
Total dietary fiber		0.45	0.61	0.51	0.64			0.56			0.36	0.36
Median	0.48	0.46	0.40	0.50	0.50	0.55	0.41	0.44	0.49	0.46	0.44	0.37
Average	0.38	0.46	0.42	0.47	0.51	0.57	0.45	0.39	0.46	0.47	0.45	0.37

* : Energy-adjusted Pearson's correlation coefficient of nutrient intakes, except for energy, calculated with FFQ and DRs.[25]

† : De-attenuated, log-transformed and energy-adjusted Pearson's correlation coefficient of nutrient intakes, except for energy, measured with FFQ and DRs.[24]

‡ : De-attenuated and energy-adjusted Pearson's correlation coefficient of nutrient intakes, except for energy, calculated with FFQ and DRs.[26]

§ : Energy-adjusted Pearson's correlation coefficient of nutrient intakes, except for energy, computed with FFQ and DRs.[27]

∣ : De-attenuated, age- and energy-adjusted Spearman's rank correlation coefficients of nutrient intakes, except for energy, measured with FFQ and DRs.[28]

WDR: weighed diet record, FFQ: food frequency questionnaire, DR: diet record.

- 83.0), and 21.7 ± 2.2 (16.9 - 28.4) for females, respectively.

Intake of nutrients

The intakes of energy and macro- and micro-nutrients gauged with the FFQ were generally lower than those with 3d-WDRs (Table 1). The ratios of nutrient consumption measured with the FFQ vs. 3d-WDRs (minimum- median- maximum) were distributed from 0.57 - 0.79 - 1.09 for males and 0.61 - 0.86 -1.04 for females.

De-attenuated, log-transformed and energy-adjusted Pearson's CCs between intakes of nutrients quantified with the FFQ and 3d-WDRs were distributed from 0.12 (n-6 PUFAs) - 0.45 (vitamin C) - 0.86 (carbohydrate and carbohydrate energy %) for males (Table 2), and energy-adjusted Spearman's rank CCs were distributed from 0.13 (n-6 PUFAs) - 0.35 (protein energy % and vitamin D) - 0.76 (carbohydrate energy %).

De-attenuated, log-transformed and energy-adjusted Pearson's CCs between intakes of nutrients quantified with the FFQ and 3d-WDRs were distributed from 0.10 (vitamin B_1) - 0.38 (carotene and folate) - 0.66 (carbohydrate energy %) for females (Table 3), and energy-adjusted Spearman's rank CCs were distributed from 0.11 (vitamin B_1) - 0.34 (protein energy % and SFAs) - 0.47 (calcium).

Median percentages of exact agreement, agreement within adjacent categories, and disagreement according to the quartile classification of energy-adjusted nutrient intakes quantified with the FFQ and 3d-WDRs were 33, 74, and 5 for males (Table 4), and 35, 76, and 7 for females (Table 5), respectively.

DISCUSSION

Because our FFQ is brief, covering 47 foods/food groups, the mean daily intakes of energy and 26 macro- and micro-nutrients determined with the FFQ were, as expected, generally smaller than those measured with the 3d-WDRs.[19-21] De-attenuated, log-transformed and energy-adjusted Pearson's CCs between intakes of selected nutrients quantified with the FFQ and 3d-WDRs were distributed from 0.10 - 0.86 and energy-adjusted Spearman's rank CCs were from 0.11 to 0.76. For most nutrients, fairly high relative validity values for the FFQ were achieved with reference to the 3d-WDRs. But the disagreement values for certain nutrients were not negligible and non-differential misclassification will unduly underestimate the risk.[22] Our FFQ thus should be deliberately applicable to rank individuals according to consumption of energy and nutrients selected for dietary studies in middle-aged Japanese.

Relative validity values are dependent on various parameters, such as person, place, time, and study protocols, which include the study subjects (e.g., people in the general population vs. dietitians/nurses), study devices adopted, interval between the two batteries of tests studied, sequence of the batteries, number of food items in the FFQ, procedures and days of DRs, and diversity of food intake (e.g. Japanese, Chinese and American diets).[2-4]

Relative validity values for macronutrients and respective energy % were reasonably high, but those for some micronutrients, including cholesterol, vitamins, minerals and dietary fibers, were rather low because the two methods measured different profiles of dietary consumption. The former inquired about dietary habits during the preceding year, and the latter surveyed actual food intakes for 3 days. WDRs are accurate without recall bias, but do not necessarily indicate habitual food consumption. Naturally, the two values do not necessarily correlate well with each other. It is also well known that great intra-individual variation exists by day, week and season for micronutrients, including vitamins and minerals.[16-19, 23] Three days are not long enough to assess the actual consumption of those nutrients and relative validity indices are invariably low, particularly for nutrients with high within-individual variation. Thus, short-day WDRs cannot be accepted as the gold standard. Furthermore, the both values estimated with FFQ and 3d-WDRs appear underestimated partly because incompleteness of the database published.[9-12] Accordingly, our investigation should rightly be called a "*relative*" validation study, and the indices need to be carefully evaluated.

An FFQ covering 47 foods/food groups may not be adequate for accurately assessing consumption of energy and 26 macro- and micro-nutrients. We formerly developed an SQFFQ with 118 foods/food groups. Its relative validity indices against 28d-WDRs (consecutive 7 day-WDRs x 4 seasons) were more favorable than with the short FFQ,[6] which may be partly explained by the number of included foods/food groups. In general, the greater the number of foods/food groups listed in the FFQ, the higher the relative validity values, but the lower the compliance among study subjects.[2-4] In addition, the fact that portion/serving size is requested by the SQFFQ, but not by the FFQ, except for staple foods, may be another reason for variation in the relative validity indices.

Because our long SQFFQ was applied to Japanese dietitians,[6] it is also plausible that the relative validity indices were more favorable than with subjects from the general populace. Reducing the study subjects' burden appears critical and questionnaires should be designed to be reasonably short when self-administered by the general public, especially for large-scale epidemiological studies. We thus had to shorten our questionnaire to maintain high compliance and still be able to rank the study subjects according to their nutrient intakes.

The sequence of application of study devices also appears crucial,[2-4] The FFQ should be first administered and relative validity figures then evaluated with DRs/WDRs distributed later because FFQs are delivered to the study subjects in the actual dietary epidemiology settings. With the reverse order, DRs/WDRs invariably yield education/memory effects, by which relative validity values are artificially improved, particularly when the interval between the two batteries of tests is short.

Here, we compared the relative validity values for a short FFQ with less than 100 food items applied to the Japanese general populace. Egami et al.[24] earlier administered a 97-item FFQ before

WDRs (Table 6) and their relative validity indices were almost equivalent to those of our questionnaire, with values for macronutrients also consistently greater than those for micronutrients, including vitamins and minerals. Other DRs/WDRs were delivered prior to respective FFQs,[25-28] but as discussed earlier, the figures should be carefully interpreted. The relative validity values for most nutrients in our questionnaire nonetheless stand comparison not only with Japanese data but also with those for brief FFQs employed elsewhere in the world.[19,21]

In conclusion, relative validity values were rather low for several nutrients, but satisfactorily high figures were obtained with most nutrients for our FFQ against the 3d-WDRs values. The questionnaire thus seems applicable to rank individuals according to consumption of energy and nutrients selected in dietary studies in the middle-aged Japanese. Bearing in mind these strengths and weaknesses of our FFQ, it can be administered to the general populace, with caution, to investigate possible associations between dietary intake and disease/health in case-control and cohort studies.

ACKNOWLEDGMENTS

The authors thank the volunteers for their participation in the present study, and express their appreciation to Ms. Y. Miyai and Ms. M. Sato for their technical assistance in preparing this manuscript.

REFERENCES

1. World Cancer Research Fund/American Institute for Cancer Research. Food, Nutrition and the Prevention of Cancer: a Global Perspective. Washington, DC: American Institute for Cancer Research, 1997.
2. Thompson FE, Byers T. Dietary assessment resource manual. J Nutr 1994; 124: S2245-S2317.
3. Willett W. Nutritional Epidemiology, 2nd Edition. New York: Oxford University Press, 1998.
4. Margetts BM, Nelson M. Design Concepts in Nutritional Epidemiology. Oxford: Oxford University Press, 1999.
5. Tokudome S, Ikeda M, Tokudome Y, Imaeda N, Kitagawa I, Fujiwara N. Development of data-based semi-quantitative food frequency questionnaire for dietary studies in middle-aged Japanese. Jpn J Clin Oncol 1998; 28: 679-87.
6. Tokudome S, Imaeda N, Tokudome Y, Fujiwara N, Nagaya T, Sato J, et al. Relative validity of a semi-quantitative food frequency questionnaire versus 28 day weighed diet records in Japanese female dietitians. Eur J Clin Nutr 2001; 55: 735-42.
7. Imaeda N, Fujiwara N, Tokudome Y, Ikeda M, Kuriki K, Nagaya T, et al. Reproducibility of a semi-quantitative food frequency questionnaire in Japanese female dietitians. J Epidemiol 2002; 12: 45-53.
8. Tokudome S, Goto C, Imaeda N, Tokudome Y, Ikeda M, Maki S. Development of a data-based short food frequency questionnaire for assessing nutrient intake by middle-aged Japanese. Asian Pacific J Cancer Prev 2004; 5: 40-3.
9. Science and Technology Agency, Japan. Standard Tables of Food Composition in Japan. The Fourth Edition. Tokyo: Ministry of Finance, 1982. (in Japanese)
10. Science and Technology Agency, Fatty acids, Cholesterol, Vitamin E Composition Table of Japanese Foods. Tokyo: Ishiyaku Shuppan, 1989. (in Japanese)
11. Science and Technology Agency, Japan. Standard Tables of Food Composition in Japan. The Fifth Edition. Tokyo: Ministry of Finance, 1993. (in Japanese)
12. Science and Technology Agency, Japan. Follow-up of Standard Tables of Food Composition in Japan. Tokyo: Ishiyaku Shuppan, 1992. (in Japanese)
13. Imaeda N, Tokudome Y, Fujiwara N, Nagaya T, Kamae M, Tsunekawa S. et al. Data checking and standardization in a weighed food dietary record survey. Jpn J Nutr 2000; 58: 67-76. (in Japanese)
14. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. Am J Epidemiol 1986; 124: 17-27.
15. Liu K, Stamler J, Dyer A, McKeever J, McKeever P. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. J Chron Dis 1978; 31: 399-418.
16. Beaton GH, Milner J, Corey P, McGuire V, Cousins M, Stewart E, et al. Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Am J Clin Nutr 1979; 32: 2456-9.
17. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. Am J Epidemiol 1988; 127: 377-86.
18. Thompson RL, Margetts BM. Comparison of a food frequency questionnaire with a 10-day weighed record in cigarette smokers. Int J Epidemiol 1993; 22: 824-33.
19. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985; 122: 51-65.
20. Yarnell JWG, Fehily AM, Milbank JE, Sweetnam PM, Walker CL. A short dietary questionnaire for use in an epidemiological survey: comparison with weighed dietary records. Human Nutr Applied Nutr 1983; 37A: 103-12.
21. Johansson I, Hallmans G, Wikman A, Biessy C, Riboli E, Kaaks R. Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. Public Health Nutr 2002; 5: 487-96.
22. Rothman KJ, Greenland S. Modern Epidemiology, 2nd ed. Philadelphia: Lippincot-Raven, 1998.
23. Tokudome Y, Imaeda N, Nagaya T, Ikeda M, Fujiwara N,

Sato J, et al. Daily, weekly, seasonal, within- and between-individual variation in nutrient intake according to four season consecutive 7 day weighed diet records in Japanese female dietitians. J Epidemiol 2002; 12: 85-92.
24. Egami I, Wakai K, Kato K, Lin Y, Kawamura T, Tamakoshi A, et al. A simple food frequency questionnaire for Japanese diet: Part II. Reproducibility and validity for nutrient intakes. J Epidemiol 1999; 9: 227-34.
25. Takatsuka N, Kawakami N, Kawai K, Okamoto Y, Ishiwata K, Shimizu H. Validation of recalled food intake in the past in a Japanese population. J Epidemiol 1996; 6: 9-13.
26. Tsubono Y, Kobayashi M, Sasaki S, Tsugane S, JPHC. Validity and reproducibility of a self-administered food frequency questionnaire used in the baseline survey of the JPHC Study Cohort I. J Epidemiol 2003; 13 (Suppl I): S125-33.
27. Lee KY, Uchida K, Shirota T, Kono S. Validity of a self-administered food frequency questionnaire against 7-day dietary records in four seasons. J Nutr Sci Vitaminol 2002; 48: 467-76.
28. Ogawa K, Tsubono Y, Nishino T, Watanabe Y, Ohkubo T, Watanabe T, et al. Validation of a food-frequency questionnaire for cohort studies in rural Japan. Public Health Nutr 2003; 6: 147-57.

British Journal of Cancer (2005) 93, 15–22

www.bjcancer.com

Soybean products and reduction of breast cancer risk: a case–control study in Japan

K Hirose*[,1], N Imaeda[2], Y Tokudome[3], C Goto[3], K Wakai[1], K Matsuo[1], H Ito[1], T Toyama[4], H Iwata[4], S Tokudome[5] and K Tajima[1]

[1]*Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, 1-1 Kanokoden Chikusa-ku, Nagoya 464-8681, Japan;* [2]*Nagoya Women's University, Mizuho-ku, Nagoya 467-8610, Japan;* [3]*Nagoya Bunri University, Inazawa 492-8520, Japan;* [4]*Department of Breast Surgery, Aichi Cancer Center Hospital, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan;* [5]*Department of Health Promotion and Preventive Medicine, Nagoya City University Graduate School of Medical Science, Mizuho-ku, Nagoya 467-8601, Japan*

Components of the Japanese diet, which might contribute to the relatively low breast cancer incidence rates in Japan, have not been clarified in detail. Since soybean products are widely consumed in Japan, a case–control study taking account of the menopausal status was conducted using data from the hospital-based epidemiologic research program at Aichi Cancer Center (HERPACC). In total, 167 breast cancer cases were included and 854 women confirmed as free of cancer were recruited as the control group. Odds ratios (OR) and 95% confidence intervals (95% CI) were determined by multiple logistic regression analysis. There were reductions in risk of breast cancer associated with high intake of soybean products among premenopausal women. Compared with women in the lowest tertile, the adjusted ORs for top tertile intake of *tofu* (soybean curd) was 0.49 (95% CI, 0.25–0.95). A significant decrease in premenopausal breast cancer risk was also observed for increasing consumption of isoflavones (OR = 0.44; 95% CI, 0.22–0.89 for highest *vs* lowest tertile; *P* for trend = 0.02). The present study found a statistically inverse association between *tofu* or isoflavone intake and risk of breast cancer in Japanese premenopausal women, while no statistically significant association was evident with the risk among postmenopausal women.

British Journal of Cancer (2005) **93,** 15–22. doi:10.1038/sj.bjc.6602659 www.bjcancer.com
Published online 7 June 2005

Keywords: soybean products; isoflavones; menopausal status

Despite a marked increase in recent years, the incidence rates for female breast cancer in most Asian countries are much lower than those in the Western world (Parkin *et al*, 1997). Much of the international variation is due to differences in established reproductive risk factors, such as age at menarche, parity and age at first birth, but dietary habits might also contribute. Dietary studies of breast cancer have typically focused on the hypothesis that there is a positive link with intake of fat. While ecological studies have suggested associations in terms of incidence and mortality, leading prospective studies of breast cancer, including the Nurse's Health Study (Stampfer *et al*, 1987) and studies of large cohorts from New York (Toniolo *et al*, 1994) and Norway (Vatten *et al*, 1990) have shown no relation (Willett, 1997). Recently, there has been more interest in other dietary factors, such as soybean products, which may protect against breast cancer and provide an explanation for some of the international differences in incidence rates (Adlercreutz, 1990; Messina and Barnes, 1991).

Soybeans provide a unique concentrated source of isoflavones and soybeans or isoflavones have been shown to exert anti-carcinogenic effects on hormone-related cancers in a large number of experimental studies. Despite the growing interest in the protective effects, there are relatively few epidemiological data available since soybean products are consumed mainly by Asian populations. In Japan, intake is in various forms, including *tofu* (soybean curd), *okara* (*tofu* lees), *moyashi* (soybean sprouts), *tonyu* (soymilk), *yuba* (soy milk skin), *kinako* (soy flour), *miso* (fermented soybean paste), *atsuage* (deep fried *tofu*), *aburage* (thinly sliced deep fried *tofu*), *natto* (fermented soybeans), *koyadofu* (freeze dried *tofu*) and *shoyu* (soy sauce), so that the diet is likely to be much richer in isoflavones than in the Western world, where the diet usually does not include soybeans. We here evaluated the association between risk of breast cancer and consumption of soybean products and isoflavones using data from the hospital-based epidemiologic research program at Aichi Cancer Center (HERPACC).

*Correspondence: Dr K Hirose; E-mail: khirose@aichi-cc.jp
Received 21 February 2005; revised 9 May 2005; accepted 9 May 2005; published online 7 June 2005

MATERIALS AND METHODS

Data collection

Details of the study design and subject characteristics have been described elsewhere (Yoo *et al*, 1992; Hirose *et al*, 1995, 1999, 2001). In brief, we have conducted the HERPACC study since 1988, whereby questionnaire survey is completed by first-visit out-patients to the Aichi Cancer Center Hospital (ACCH) (Tajima *et al*, 2000). All questionnaires are then collected after checking for incomplete responses by a trained interviewer and the data are loaded into the computer system of the Aichi Cancer Center Research Institute. The data collected are linked with the

hospital-based cancer registry files. This study was approved by the Institutional Review Board, and all participants provided written, informed consent.

The questionnaire included questions on occupation, medical history, height, weight, weight at around 20 years of age, family history (parents and siblings), smoking and drinking habits, sleeping habits, physical exercise and reproductive history. The details were taken prior to assessment of symptoms and all information was collected before clinical diagnoses were made.

Cases and controls

The present analysis was restricted to women aged 30 and over who visited hospitals between January 2001 and March 2002. As eligible cases, 79 premenopausal and 88 postmenopausal women diagnosed as having breast cancer by histological examination within 6 months of the first-visit were recruited. A methodological study applying the same HERPACC data set earlier showed that the OR based on a large number of controls gives more power and steadier estimates than the use of matched controls (Hamajima *et al*, 1994); therefore, we used all noncancer individuals as candidates for controls in this study. As the controls, 854 female first-visit outpatients who had never been diagnosed as having cancer were recruited. Table 1 summarises details for the 167 cases and 854 controls by age group and selected characteristics.

Exposure data

The interview included a validated, semiquantitative food-frequency questionnaire (FFQ) (Tokudome *et al*, 1998, 2001; Imaeda *et al*, 2002). Dietary intake was ascertained using a detailed quantitative FFQ, including 119 food items/recipes, and the following food groups were included: (a) meat and meat products (poultry, ground meat, pork meat, beef, ham, sausage, bacon, liver); (b) fish and fish products (salmon, eel, pale blue fleshed fish, red fleshed fish, white fleshed fish, squid, octopus, shrimp, crab, dried fish larvae, small bony fish, tuna canned in oil, cod roe, oyster, shellfish, dried squid, fried fish paste, fish paste sausage); (c) green-yellow vegetables (green leafy vegetables, pumpkin, carrot, broccoli, green pepper, green soybean, green beans, tomato); (d) other vegetables (cabbage, Japanese radish, burdock, bamboo shoots, cucumber, eggplant, lettuce, bean sprout, onion, Chinese cabbage, Japanese radish); (e) fruits (oranges, mandarin oranges, persimmon, banana, apple, strawberries, kiwi, peach, grapes, watermelon, melon, Japanese pears); (f) dairy products (low fat milk, medium fat milk, high fat milk, calcium enriched milk and yoghurt, skim milk, lactic acid bacteria beverage, yoghurt, cheese, ice cream).

Furthermore, the following soybean products were included: (a) *tofu* (soybean curd); (b) *miso* (fermented soybean paste); (c) *atsuage* (deep fried *tofu*); (d) *aburage* (thinly sliced deep fried *tofu*); (e) *natto* (fermented soybeans); (f) *koyadofu* (freeze dried *tofu*). We evaluated validity of intakes based on the questionnaire against those according to 28-day (four-season consecutive 7-day) weighted diet records among 79 Japanese female dietitians. In the validation study, the Spearman's correlation coefficients between the estimate intake of soybean products and isoflavone from the questionnaire and that from dietary records were 0.51, 0.53, respectively. For reproducibility of estimation from the questionnaire, the Spearman's correlation coefficients for the consumption of soybean products and isoflavone intake between two questionnaires administered 1 year apart were 0.57 and 0.47, respectively.

All subjects in the present study were asked for average frequency and portion size of consumption, during the period of 1 year before onset of the present disease or before the interview. There were eight categories of possible responses, ranging from 'rarely or never' to 'three or more times per day'. For each food, a commonly used unit or portion size was specified and the interviewers asked the subjects using sampling models of full-scale photographs. We ascertained average amount of daily consumption of each food and nutrients by multiplying the food intake (in grams) or serving size and the nutrient content per 100 grams of food as listed in the Standard Tables of Food Composition, Version 5 and the Follow-up of Standard Tables of Food Composition (Science and Technology Agency, Japan, 1994, 2001). Isoflavone intakes were separately calculated from USDA-Iowa State University Database on the Isoflavone Content of

Table 1 Basic characteristics of cases and controls

	Cases ($n=167$)	Controls ($n=854$)
Age (years)		
30–39	19 (11.4%)	99 (11.6%)
40–49	46 (27.5%)	279 (32.7%)
50–59	54 (32.3%)	280 (32.8%)
⩾60	48 (28.7%)	196 (23.0%)
Mean age (s.d.)	52.7 (10.2)	51.4 (10.5)
Motives for consultation		
Self-recommendation	47 (28.1%)	286 (33.5%)
Family recommendation	43 (25.8%)	194 (22.7%)
Referral from other clinics	44 (26.4%)	139 (16.3%)
Secondary screening after primary screening	33 (19.8%)	224 (26.2%)
Others	0 (0.0%)	6 (0.7%)
Unknown	0 (0.0%)	5 (0.6%)
Smoking status		
Never	146 (87.4%)	710 (83.2%)
Ever	9 (5.4%)	51 (6.0%)
Current	12 (7.2%)	91 (10.7%)
Unknown	0 (0.0%)	1 (0.1%)
Drinking status		
Never	111 (66.5%)	546 (63.9%)
Ever	0 (0.0%)	13 (1.5%)
Current	56 (33.5%)	295 (34.5%)
Exercise		
No	54 (32.3%)	291 (34.1%)
⩽60 min/week	33 (19.8%)	192 (22.5%)
⩽120 min/week	27 (16.2%)	126 (14.8%)
>120 min/week	47 (28.1%)	225 (26.4%)
Unknown	6 (3.6%)	20 (2.3%)
Mean BMI (s.d.)	22.9 (3.1)	22.0 (3.0)
Mean of age at first birth (s.d.)	25.8 (3.6)	25.7 (3.5)
Mean of age at menarche (s.d.)	13.5 (1.5)	13.4 (1.6)
Parity		
Parous	155 (93.4%)	763 (89.7%)
Nulliparous	11 (6.6%)	88 (10.3%)
Menopausal status		
Premenopausal	79 (47.3%)	414 (48.5%)
Postmenopausal	88 (52.7%)	440 (51.5%)
Age at menopause among postmenopausal women		
⩽47	24 (27.3%)	105 (23.9%)
48–52	42 (47.7%)	225 (51.1%)
⩾53	22 (25.0%)	83 (18.9%)
Unknown	0 (0.0%)	27 (6.1%)
Family history[a]		
No	153 (91.6%)	791 (92.6%)
Yes	14 (8.4%)	63 (7.4%)

s.d. = standard deviation. [a]Family history in mother or sisters.

Foods, Release 1.3-2002 (http://www.nal.usda.gov/fnic/foodcopm/Data/isoflav/isoflav.html). Isoflavone intake was calculated using consumption of six items of soybean products, green soybean, peanuts, Japanese green tea, and vegetables other than green-yellow vegetables such as cucumber, eggplant, lettuce, bean sprouts, onion, Chinese cabbage.

Statistical analysis

Dietary intake data were analysed by individual food items, food groups and nutrients for all subjects combined and separately for premenopausal and postmenopausal women. The differences of means were examined by *t*-test and all *P*-values presented are two-sided.

Logistic regression analysis was used to obtain odds ratios (ORs) and 95% confidence intervals (95% CI) as estimates of relative risk. The *P*-value for trend corresponded to the estimate of the slope derived from the logistic model in the case that the integers, 1 to *n*, were assigned to the ordered *n* levels of each factors. The LOGISTIC procedure provided by SAS (SAS Institute, Cary, NC, USA) was utilised to perform the calculations. To compare differences and similarities of effects of Japanese diet on risk for breast cancer, subjects were stratified with reference to menopausal status. The ORs of breast cancer are calculated for the selected food groups and isoflavone intake in tertiles, with the lowest tertile as the referents. Multivariate models was adjusted for age, energy (as a continuous variable), motives for consultation (self-recommendation, family recommendation, referral from other clinics, secondary screening after primary screening, others), smoking status (never, ever, current), drinking status (never, ever, current), exercise (none, $\leqslant$60 min/week, $\leqslant$120 min/week, $>$120 min/week), family history (yes or no), age at menarche ($\leqslant$12, 13, $\geqslant$14), parity (0,1,2, 3+), age at first full-term pregnancy ($\leqslant$23, 24–27, $\geqslant$28), and current body mass index (BMI) ($\leqslant$20, 20–25, $\geqslant$25). Further, inclusion in the model of age at menopause ($\leqslant$47, 48–52, $\geqslant$53) was performed with the calculations for postmenopausal women. Energy was adjusted for the multivariate nutrient density method.

A positive family history may involve different types and numbers of relatives. Distant relatives share less genetic influence and fewer confounding environmental and/or behavioral factors than do close ones. Furthermore, information on the medical history of distant relatives is limited and less precise than that of close or first-degree relatives. In this study, the presence of either a mother or sister with breast cancer was considered as a positive family history. BMI was calculated as weight/height^2 (kg m^{-2}), according to Quetelet's formula and current BMI values were stratified into three categories. Since a BMI$\geqslant$25 is defined as obese by the Japanese Society for the Study of Obesity, the cutoff for the highest BMI group was BMI$\geqslant$25.

RESULTS

Table 2 summarises data for daily intake of some nutrients and consumption of selected food groups among cases and controls. There were no significant differences between cases and controls in intake of energy and fat. The means of meat and meat products intakes per day was 60.7 g (standard deviation (s.d.), 34.7 g) and 67.0 g (s.d., 39.4 g), respectively, for cases and controls ($P<0.04$). Among premenopausal women, the means of soybean products were 51.7 g (s.d., 31.2 g) in case group and 63.5 g (s.d., 38.8 g) in control group ($P<0.01$, data not shown). Breast cancer cases reported lower total isoflavones intake per day than controls, with averages of 20.8 mg (s.d., 10.8 mg) and 25.8 mg (s.d., 14.3 mg), respectively ($P<0.0001$, data not shown) among premenopausal women. We observed no other significant differences among premenopausal women and there were no significant differences

Table 2 Distribution of selected dietary variables (intake/day) among breast cancer cases and controls

	Cases				Controls			
		Percentile				Percentile		
Nutritional factors and food items	**Mean**	**25**	**50**	**75**	**Mean**	**25**	**50**	**75**
Nutritional factors								
Energy (kcal)	1831	1568	1791	2038	1859	1593	1835	2093
Total protein (g)	72.1	59.0	70.1	83.0	73.7	60.8	72.1	83.7
Total fat (g)	57.5	43.5	55.0	70.6	59.1	45.5	57.9	70.4
% of energy (%)	27.9	25.0	28.7	31.4	28.4	24.3	28.5	32.3
Carbohydrate (g)	252.9	220.8	243.6	279.7	254.7	217.5	249.8	289.7
Total dietary fibre (g)	14.4	11.4	13.6	16.7	14.8	11.5	14.2	17.2
Vitamin C (mg)	161.5	115.6	152.1	194.4	164.9	108.4	146.3	199.3
Vitamin E (mg)	10.1	7.9	9.7	12.1	10.3	8.1	10.0	12.1
Isoflavones (mg)	24.8	15.3	21.3	30.3	27.1	17.0	24.2	32.5
Food items								
Meat and meat products (g)	60.7	36.4	55.7	77.9	67.0	38.6	61.6	90.0
Fish and fish products (g)	56.5	38.9	52.9	68.6	56.8	36.0	52.9	71.4
Green-yellow vegetables (g)	136.5	78.2	118.9	173.2	144.6	78.2	118.0	180.4
Other vegetables (g)	130.7	82.5	115.7	154.3	135.8	88.2	125.4	172.1
Fruit (g)	164.5	83.4	142.1	217.0	167.1	84.1	143.3	219.7
Dairy products (g)	200.2	87.1	189.3	274.3	199.0	96.8	188.9	265.0
Soybean products (g)	63.4	35.5	53.5	82.4	67.7	39.5	60.5	84.3
Tofu (soybean curd) (g)	39.0	20.0	29.0	50.0	41.7	23.0	35.0	52.0
Miso (fermented soybean paste) (g)	6.3	2.0	5.0	10.0	7.1	5.0	8.0	10.0
Atsuage (deep fried tofu) (g)	7.4	2.5	5.0	10.0	7.3	2.5	5.0	10.0
Aburage (thinly sliced deep fried tofu) (g)	1.7	0.6	1.2	3.0	1.7	0.6	1.2	2.4
Natto (fermented soybeans) (g)	9.2	4.0	7.2	10.0	10.3	4.0	8.0	16.0
Koyadofu (freeze dried tofu) (g)	1.9	0.0	0.0	4.0	2.0	0.0	0.0	4.0

between cases and controls in average intake of food groups and nutrients among postmenopausal women.

Among premenopausal women, breast cancer risk was inversely associated with consumption of soybean products (Table 3). The ORs was 0.53 (95% CI, 0.27–1.04) for the top tertile of soybean product intake compared with the lowest tertile of intake (trend test, $P=0.06$). Among postmenopausal women, on the other hand, consumptions of soybean products, meat and meat products, fish and fish products, vegetables, fruits, dairy products were not associated with the risk of breast cancer.

The ORs for breast cancer according to type of soybean product are presented in Table 4. A statistically significant inverse relation was observed between *tofu* consumption and breast cancer in premenopausal women. ORs were 0.44 (95% CI, 0.22–0.90), 0.49 (95% CI, 0.25–0.95) for the second to the top tertiles of intake compared with the lowest tertile of intake of *tofu* (trend test, $P=0.03$). Compared with those in the lowest tertile of *atsuage* (deep fried tofu) consumption, the adjusted OR for breast cancer in top tertile was 0.70 (95% CI, 0.35–1.38) among premenopausal women. On the other hand, a significantly increased risk of breast cancer with consumption of *atsuage* was observed in postmenopausal women. The adjusted OR for top tertile of *atsuage* consumption was 2.28 (95% CI, 1.15–4.51, trend test $P=0.02$). Also, consumption of *aburage* (thinly sliced deep fried tofu) showed a similar positive association with the risk of breast cancer. There were no associations with intake of *tofu*, *miso* and *natto* intake among postmenopausal women. With intake of *koyadofu* (freeze dried tofu) divided into tertiles, there was no apparent modification of the breast cancer risk among either pre- or postmenopausal women.

We next determined the association between total intake of isoflavone and risk of breast cancer and found a statistically significant inverse association among premenopausal women. Compared with women in the lowest tertile of isoflavone consumption, the top tertile had an adjusted OR of 0.44 (95% CI, 0.22–0.89; trend test $P=0.02$). Among postmenopausal women the similar trend was also observed; however, there was no statistically significant association between consumption of isoflavone and breast cancer risk (Table 5).

DISCUSSION

The present study found a statistically inverse association between risk of breast cancer and soybean products (*tofu*) or isoflavones intake in Japanese premenopausal women while there was no statistically significant link among their postmenopausal counterparts.

Table 3 Odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer according to food group intake by menopausal status

	Premenopausal women				Postmenopausal women			
	OR for tertiles of food group intake (g/1000 kcal)				OR for tertiles of food group intake (g/1000 kcal)			
	1	2	3	P-value test	1	2	3	P-value test
Meat and meat products								
Tertile median	24.3	38.4	54.4		13.7	29.4	45.7	
Cases/controls	29/138	31/137	19/139		31/146	39/147	18/147	
OR[a]	1.00	1.02	0.54	0.11	1.00	1.11	0.64	0.24
(95% CI)		(0.54–1.92)	(0.26–1.12)			(0.62–2.02)	(0.32–1.28)	
Fish and fish products								
Tertile median	14.8	25.6	38.5		18.2	30.6	45.4	
Cases/controls	21/138	29/138	29/138		26/145	37/147	25/148	
OR[a]	1.00	1.53	1.36	0.46	1.00	1.42	0.77	0.42
(95% CI)		(0.74–3.18)	(0.65–2.88)			(0.75–2.66)	(0.39–1.52)	
Vegetables[b]								
Tertile median	80.4	125.3	189.6		98.6	145.2	220.5	
cases/controls	22/138	39/138	18/138		31/146	30/146	27/148	
OR[a]	1.00	1.45	0.70	0.38	1.00	1.01	0.85	0.63
(95% CI)		(0.76–2.80)	(0.32–1.50)			(0.54–1.89)	(0.44–1.63)	
Fruit								
Tertile median	27.5	62.3	115.4		46.3	93.7	158.5	
Cases/controls	28/139	29/136	22/139		24/146	32/147	32/147	
OR[a]	1.00	1.08	0.65	0.24	1.00	1.21	1.38	0.34
(95% CI)		(0.55–2.10)	(0.32–1.33)			(0.62–2.35)	(0.71–2.69)	
Dairy products								
Tertile median	34.0	98.5	158.5		32.3	103.6	165.5	
Cases/controls	30/139	22/136	27/139		22/146	31/147	35/147	
OR[a]	1.00	0.91	0.73	0.37	1.00	1.54	1.64	0.15
(95% CI)		(0.46–1.79)	(0.37–1.45)			(0.80–2.96)	(0.84–3.20)	
Soybean products								
Tertile median	17.2	29.7	47.9		20.1	35.3	56.5	
Cases/controls	36/139	23/137	20/138		31/146	28/147	29/147	
OR[a]	1.00	0.60	0.53	0.06	1.00	0.87	0.70	0.28
(95% CI)		(0.30–1.18)	(0.27–1.04)			(0.47–1.61)	(0.37–1.33)	

[a]Adjusted for age, motives for consultation, smoking, drinking, exercise, energy, family history, age at menarche, parity, age at first full-term pregnancy, BMI and age at menopause for postmenopausal women. [b]Included both green-yellow vegetabeles and other vegetables.

Table 4 Odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer according to soybean product intake by menopausal status

	Premenopausal women				Postmenopausal women			
	OR for tertiles of soybean product intake (g/1000 kcal)				OR for tertiles of soybean product intake (g/1000 kcal)			
	1	2	3	*P*-value test	1	2	3	*P*-value test
Tofu (soybean curd)								
Tertile median	8.9	17.1	31.3		10.8	19.6	35.9	
Cases/controls	38/137	20/139	21/137		26/145	36/148	26/147	
OR[a]	1.00	0.44	0.49	0.03	1.00	1.34	0.71	0.34
(95% CI)		(0.22–0.90)	(0.25–0.95)			(0.73–2.44)	(0.36–1.39)	
Miso (fermented soybean paste)								
Tertile median	1.4	3.4	5.5		1.5	3.9	5.8	
Cases/controls	30/137	30/137	19/136		37/145	23/147	28/147	
OR[a]	1.00	1.14	0.58	0.15	1.00	0.52	0.64	0.11
(95% CI)		(0.59–2.19)	(0.28–1.20)			(0.27–0.98)	(0.34–1.17)	
Atsuage (deep fried tofu)								
Tertile median	0.0	2.9	5.9		0.0	3.1	6.4	
Cases/controls	29/139	24/137	26/138		20/146	33/146	35/148	
OR[a]	1.00	0.67	0.70	0.31	1.00	1.95	2.28	0.02
(95% CI)		(0.34–1.34)	(0.35–1.38)			(0.98–3.86)	(1.15–4.51)	
Aburage (thinly sliced deep fried tofu)								
Tertile median	0.0	0.5	1.4		0.3	0.7	1.9	
Cases/controls	23/136	34/136	22/137		22/145	31/147	35/147	
OR[a]	1.00	1.67	1.07	0.97	1.00	1.75	1.62	0.17
(95% CI)		(0.82–3.40)	(0.51–2.26)			(0.89–3.43)	(0.83–3.14)	
Natto (fermented soybeans)								
Tertile median	1.2	3.7	8.7		1.3	4.0	10.8	
Cases/controls	29/138	26/137	24/137		30/145	30/145	27/148	
OR[a]	1.00	0.89	0.84	0.56	1.00	1.00	0.79	0.47
(95% CI)		(0.46–1.74)	(0.43–1.64)			(0.54–1.87)	(0.41–1.51)	
Koyadofu (freeze dried tofu)								
Tertile median	0.0	1.9	3.6		0.0	2.1	4.2	
Cases/controls	55/297	6/57	18/60		58/284	14/77	16/79	
OR[a]	1.00	0.68	1.38	0.52	1.00	0.81	0.99	0.84
(95% CI)		(0.27–1.73)	(0.69–2.79)			(0.40–1.64)	(0.50–1.97)	

[a]Adjusted for age, motives for consultation, smoking, drinking, exercise, energy, family history, age at menarche, parity, age at first full-term pregnancy, BMI and age at menopause for postmenopausal women.

Soybean foods are rich in precursors of the isoflavone daidzein and genistein, which are heterocyclic phenols similar in structure to oestrogens, and it has been hypothesised that a high dietary intake of soybean products might reduce breast cancer risk by interfering with the action of endogenous oestradiol (Messina, 1999). The results are in line with the inverse association between intake of soybean products and breast cancer risk suggested from ecological/cross-sectional studies (Adlercreutz, 1995; Adlercreutz and Mazur, 1997), and also from analytical investigations. Thus, case–control studies have found that soybean food intake was associated with a decreased risk of breast cancer among premenopausal Singapore women (Lee *et al*, 1991), and both pre- and postmenopausal Asian-American women (Wu *et al*, 1996), although a Chinese case–control study failed to detect any protective effects of soybean food (Yuan *et al*, 1995). Cohort studies among Japanese (Hirayama, 1990; Yamamoto *et al*, 2003), Japanese-American (Nomura *et al*, 1978) and Caucasian-American women (Greenstein *et al*, 1996) have also provided some evidence that soybean products may reduce the risk of breast cancer. A prospective study conducted in Japan, however, found no link between soya consumption and breast cancer risk, but in this case the subjects were city residents in Hiroshima or Nagasaki exposed to high doses of ionising radiation and therefore the cohort was unusual (Key *et al*, 1999). A recent cohort study based on public health center in Japan (Yamamoto *et al*, 2003) found frequent *miso soup* and isoflavone consumption to be associated with a reduced risk of breast cancer and the protective effect was stronger in postmenopausal women. However, the FFQ applied included only two items for soybean-ingredient foods (i.e. *miso soup* and soyfoods), making it impossible to investigate differences in effects among types of soybean-ingredient foods.

Isoflavone intake by Japanese is much higher than that by Western populations (Jones *et al*, 1989; Kimira *et al*, 1998). *Tofu*, *miso* and *natto* were main foods containing rich isoflavones. Attributable rate of genistein were *tofu* (49.6%), *miso* (20.9%), *natto* (14.7%) among Japanese. In the present study, *tofu* was protective in premenopausal women, while *atsuage and aburage*, deep fried *tofu* containing much oil, was associated with elevated risk among postmenopausal women. This may be due to fat intake, which can exert an influence on the development of breast cancer among postmenopausal women. Some studies have suggested that high intake of soybean products in premenopausal women may

Table 5 Odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer according to isoflavone intake by menopausal status

	Premenopausal women				Postmenopausal women			
	OR for tertiles of isoflavone intake (mg/1000 kcal)				OR for tertiles of isoflavone intake (mg/1000 kcal)			
	1	2	3	*P*-value test	1	2	3	*P*-value test
Tertile median	7.61	11.87	18.47		8.69	13.59	22.26	
Cases/controls	36/138	24/139	19/137		33/145	29/144	26/151	
OR[a]	1.00	0.62	0.44	0.02	1.00	0.76	0.58	0.09
(95% CI)		(0.32–1.20)	(0.22–0.89)			(0.41–1.40)	(0.30–1.10)	

[a]Adjusted for age, motives for consultation, smoking, drinking, exercise, energy, family history, age at menarche, parity, age at first full-term pregnancy, BMI and age at menopause for postmenopausal women.

Clinical Studies

reduce serum oestradiol concentrations, suppress the mid-cycle surge of gonadotropins, and perhaps increase the length of the menstrual cycle (Cassidy *et al*, 1994; Lu *et al*, 1996; Nagata *et al*, 1997).

There is increasing evidence that dietary factors may play a role in the production, metabolism, and bioavailability of sex hormones. Soy-containing diets have long been known to be typical of some ethnic groups who experience low breast cancer risk. Soybeans contain a significant amount of the isoflavones daidzein and genistein, which may exert antioestrogenic effects and protect epithelial tissue from stimulation by endogenous oestrogens. There are several possible mechanisms by which soy isoflavones specifically may modulate the risk of breast carcinoma: (1) increase of serum levels of sex hormone-binding globulin (SHBG) (Adlercreutz *et al*, 1991; Mousavi and Adlercreutz, 1993); (2) downregulation of enzymes involved in oestrogen biosynthesis, such as aromatase (Adlercreutz *et al*, 1993); (3) inhibition of 17β-hydroxysteroid dehydrogenase type I (Makela *et al*, 1995); (4) suppression of the gonadotorpins follicule stimulating hormone (FSH) and luteinising hormone (LH); (5) change in intestinal flora, which affect reabsorption of E2 and lower circulating oestrogen levels (Adlercreutz, 1998). Studies examining the effect of a soy protein diet on the menstrual cycle have demonstrated a significant increase in follicular phase length and delay in menstruation, including suppression of midcycle surges of LH and FSH, which potentially may reduce the risk of breast cancer (Cassidy *et al*, 1994; Lu *et al*, 1996; Nagata *et al*, 1998; Kumar *et al*, 2002). Isoflavones have in fact received a great deal of attention due to their antiproliferative properties and these would support the protective effect against breast cancer in premenopausal women observed in the present study. However, clarification of effects in postmenopausal women is still required. In this context it is of interest that isoflavones may exert both oestrogenic and antioestrogenic properties after modification by intestinal bacteria.

Soybeans are a unique dietary source of a group of phytochemicals and several natural anticarcinogens have now been identified in soybeans, such as protease inhibitors, phytates, phytosterols, saponins and lignans. Also, soybeans are an excellent source of dietary fibre and micronutrients, especially calcium. It will be important in future epidemiological studies to investigate the association between intake of soybean products and breast cancer by obtaining a more complete assessment of soybean intake. Further studies are required to confirm the ability of the nutritional profile of soybeans to reduce the risk of breast cancer.

Potential limitations of the present study should be considered. One methodological issue is selection of base population for controls. We applied noncancer patients at ACCH as controls because it is reasonable to assume our cases arise within this population base. Main reasons to visit at ACCH among cases and controls were self/family recommendation, referral from other clinics, and secondary screening after primary screening. Although distributions of these reasons slightly differ, it was considered in statistical analyses. Notable point of our control population is its similarity to general population in terms of exposure of interest, here dietary pattern. We have compared lifestyle characteristics between outpatients in ACCH and the 1231 individuals randomly selected from the general population, and confirmed that they are not substantially different (Inoue *et al*, 1997). Possible bias due to medical background of controls is another potential source of bias; however, our previous report revealed substantially limited impact of it. More than 66% of noncancer outpatients at ACCH did not have any specific medical condition. Remaining 34% of them have specific diseases, but common part of them were benign tumours and/or non-neoplastic polyps (13.1%), mastitis (7.5%), digestive disease (4.1%), or benign gynaecological disease (4.1%) (Hamajima *et al*, 1995). This situation is very different from that in the US, where people visit local general clinics first, and are then referred to hospitals, which function as secondary and/or specific facilities for further medical treatment. We therefore conclude that it is feasible to use noncancer outpatients as controls in epidemiological studies with due consideration of age, sex, season, and reason for visit. The present study was free of recall information bias to the questionnaire because all data were collected prior to diagnoses. Eligible controls were not matched, because our previous study showed that the large number of nonselsected controls gives a steadier estimate than selected, matched controls (Hamajima *et al*, 1994).

Another limitation of the present study included the small number of cases that are more likely to be due to chance leading to a false-positive result. Conversely, small sample may lack sufficient power to detect significant differences. The powers for detection of the OR of 0.5 for higher *vs* lower level of total isoflavone intake were 77% among premenopausal women and 81% among postmenopausal women, respectively (alpha error = 0.0500, two-sided). Much larger studies will be required to confirm the impact of soybean products and isoflavone intake on breast cancer risk among Japanese women.

A number of risk factors for breast cancer have been established, most of them related to reproductive events. Evidence from studies of migrant populations, however, has also implicated environmental or lifestyle factors as being of importance and the diet is suspected of playing a role. The present study focused on the Japanese diet, which features a high level of consumption of soybean products and was able to demonstrate that high consumption of soybean products reduces the risk in Japanese premenopausal women. The findings are biologically plausible and suggested a potential beneficial effect of soybean products and isofloavones in the prevention of breast cancer.

ACKNOWLEDGEMENTS

We thank all the doctors, nurses, technical staff and hospital business staff of Aichi Cancer Center Hospital for the daily

administration of the HERPACC study. We are greatly indebted to the staff of the Department of Breast Surgery, Aichi Cancer Center Hospital for their support and helpful discussions. We are also grateful to Ms H Fujikura, Ms K Asai, Ms K Fukaya, Ms M Obuchi, Ms C Adachi and Ms K Sanji for data collection and preparation.

This work was supported in part by a Grant-in-Aid for Cancer Research from the Japanese Ministry of Health, Labor and Welfare and a Grant-in-Aid for Scientific Research on Special Priority Areas of Cancer from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

REFERENCES

Adlercreutz H (1990) Western diet and western disease: some hormonal and biochemical mechanisms and associations. *Scand J Clin Lab Invest* 50(Suppl): 3–23

Adlercreutz H (1995) Phytoestrogens: epidemiology and a possible role in cancer protection. *Environ Health Perspect* 103: 103–112

Adlercreutz H (1998) Evolution, nutrition, intestinal microflora, and prevention of cancer: a hypothesis. *Proc Soc Exp Bio Med* 217: 241–246

Adlercreutz H, Bannwart C, Wahala K, Makeia T, Brunow G, Hase T, Arosemena PJ, Kellis J, Vickery LE (1993) Inhibition of human aromatase by mammalian lignans and isoflavonoid phytoestrogens. *J Steroid Biochem Mol Biol* 44: 147–153

Adlercreutz H, Honjo H, Higashi A, Fotsis T, Hamalainen E, Hasegawa T, Okada H (1991) Urinary excretion of lignans and isoflavonoid phytoestrogens in Japanese men and women consuming a traditional Japanese diet. *Am J Clin Nutr* 54: 1093–1100

Adlercreutz H, Mazur W (1997) Phyto-oestrogens and Western diseases. *Ann Med* 29: 95–120

Cassidy A, Bingham S, Setchell KDR (1994) Biological effects of a diet of soy protein rich in isoflavones on the menstrual cycle of premenopausal women. *Am J Clin Nutr* 60: 333–340

Greenstein J, Kushi L, Zheng W, Fee R, Campbel D, Sellers T, Folsom A (1996) Risk of breast cancer associated with intake of specific foods and food groups. *Am J Epidemiol* 143: S36

Hamajima N, Hirose K, Inoue M, Takezaki T, Kuroishi T, Tajima K (1995) Age-specific risk factors of breast cancer estimated by a case–control study in Japan. *J Epidemiol* 5: 99–105

Hamajima N, Hirose K, Inoue M, Takezaki T, Kuroishi T, Tajima K (1994) Case–control studies: matched controls or all available controls ? *J Clin Epidemiol* 47: 971–975

Hirayama T (1990) Life-style and mortality: a large-scale census-based cohort study in Japan. In *Contributions to Epidemiology and Biostatistics* Wahrendorf J (ed) Swizerland: Karger

Hirose K, Tajima K, Hamajima N, Inoue M, Takezaki T, Kuroishi K, Yoshida M, Tokudome S (1995) A large-scale, hospital-based case–control study of risk factors of breast cancer according to menopausal status. *Jpn J Cancer Res* 86: 146–154

Hirose K, Tajima K, Hamajima N, Takezaki T, Inoue M, Kuroishi T, Miura S, Tokudome S (1999) Effect of body size on breast-cancer risk among Japanese women. *Int J Cancer* 80: 349–355

Hirose K, Tajima K, Hamajima N, Takezaki T, Inoue M, Kuroishi T, Miura S, Tokudome S (2001) Association of family history and other risk factors with breast cancer risk among Japanese premenopausal and postmenopausal women. *Cancer Causes Control* 12: 349–358

Imaeda N, Fujiwara N, Tokudome Y, Ikeda M, Kuriki K, Nagaya T, Sato J, Goto C, Maki S, Tokudome S (2002) Reproducibility of a semi-quantitative food frequency questionnaire in Japanese female dietitians. *J Epidemiol* 12: 45–53

Inoue M, Tajima K, Hirose K, Hamajima N, Takezaki T, Kurosihi T, Tominaga S (1997) Epidemiological features of first-visit outpatients in Japan: comparison with general population and variation by sex, age, and season. *J Clin Epidemiol* 50: 69–77

Jones AE, Price KR, Fenwick GR (1989) Development and application of a high-performance liquid chromatographic method for the analysis of phytoestorgens. *J Sci Food Agric* 46: 357–364

Key TJ, Sharp GB, Appleby PN, Beral V, Goodman MT, Soda M, Mabuchi K (1999) Soya foods and breast cancer risk: a prospective study in Hiroshima and Nagasaki, Japan. *Br J Cancer* 81: 1248–1256

Kimira M, Arai Y, Shimoi K, Watanabe W (1998) Japanese intake of flavonoids and isoflavonoids from foods. *J Epidemiology* 8: 168–175

Kumar NG, Cantor A, Allen K, Riccardi D, Cox CE (2002) The specific role of isoflavones on estrogen metabolism in premenopausal women. *Cancer* 94: 1166–1174

Lee HP, Gourley L, Duffy SW, Esteve J, Lee J, Day NE (1991) Dietary effects on breast-cancer risk in Singapore. *Lancet* 337: 1197–1200

Lu L-JW, Anderson KE, Grady JJ, Nagamani M (1996) Effects of soya consumption for one month on steroid hormones in premenopausal women: implications for breast cancer risk reduction. *Cancer Epidemiol Biomarkers Prev* 5: 63–70

Makela S, Poutanen M, Lehtimaki J, Kostian ML, Santti R, Vihko R (1995) Estrogen-specific 17β-hydroxysteroid oxidoreductase type I (E.C.1.1.1.62) as a possible target for the action of phytoestrogens. *Proc Soc Exp Biol Med* 208: 51–59

Messina M (1999) Legumes and soybeans: overview of their nutritional profiles and health effects. *Am J Clin Nutr* 70(Suppl): 439s–450s

Messina M, Barnes S (1991) The role of soy products in reducing risk of cancer. *J Natl Cancer Inst* 83: 541–546

Mousavi Y, Adlercreutz H (1993) Genistein is an effective stimulator of sex hormone-binding globulin production in hepatocarcinoma human liver cancer cells and suppresses proliferation of these cells in culture. *Steroids* 58: 301–304

Nagata C, Kabuto M, Kurisu Y, Shimizu H (1997) Decreased serum estradiol concentration associated with high dietary intake of soy products in premenopausal Japanese women. *Nutr Cancer* 29: 228–233

Nagata C, Takatsuka N, Inaba S, Kawakami N, Shimizu H (1998) Effect of soymilk consumption on serum estrogen concentration in premenopausal Japanese women. *J Natl Cancer Inst* 90: 1830–1835

Nomura A, Henderson BE, Lee J (1978) Breast cancer and diet among the Japanese in Hawaii. *Am J Clin Nutr* 31: 2020–2025

Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J (1997) *Cancer Incidence in Five Continents* Vol. VII, IARC Scientific Publication 143. Lyon: IARC

Science and Technology Agency, Japan (1994) *Follow-up of Standard Tables of Food Composition in Japan.* Tokyo: Ishiyaku Shuppan (in Japanese)

Science and Technology Agency, Japan (2001) *Standard Tables of Food Composition in Japan* 5th edn Tokyo: Ministry of Finance (in Japanese)

Stampfer MJ, Willett WC, Colditz GA, Speizer FE (1987) Intake of cholesterol, fish and specific types of fat in relation to risk of breast cancer. In *Proceedings of the AOCS Short Course on Polyunsaturated Fatty Acids and Eicosanoids* Lands WE (ed) pp 248–252. Champaign, IL: American Oil Chemists' Society Press

Tajima K, Hirose K, Inoue M, Takezaki T, Hamajima N, Kuroishi T (2000) A model of practical cancer prevention for out-patients visiting a hospital: the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC). *Asian Pacific J Cancer Prev* 1: 35–47

Tokudome S, Ikeda M, Tokudome Y, Imaeda N, Kitagawa I, Fujiwara N (1998) Development of data-based semi-quantitative food frequency questionnaire for dietary studies in middle-aged Japanese. *Jpn J Clin Oncol* 28: 679–687

Tokudome S, Imaeda N, Tokudome Y, Fujiwara N, Nagaya T, Sato J, Kuriki K, Ikeda M, Maki S (2001) Relative validity of a semi-quantitative food frequency questionnaire versus 28 day weighted diet records in Japanese female dietitians. *Eur J Clin Nutr* 55: 732–742

Toniolo P, Riboli E, Shore RE, Pasternak BS (1994) Consumption of meat, animal products, protein, and fat and risk of breast cancer: a prospective cohort study in New York. *Epidemiology* 5: 391–397

Vatten LJ, Solvoll K, Loken EB (1990) Frequency of meat and fish intake and risk of beast cancer in a prospective study of 14,500 Norwegian women. *Int J cancer* 46: 12–15

Willett WC (1997) Specific fatty acids and risks of breast and prostate cancer: dietary intake. *Am J Clin Nutr* 66(Suppl): 1557s–1563s

Wu AH, Ziegler RG, Horn-Ross PL, Nomura AM, West DW, Kolonel LN, Rosenthal JF, Hoover RN, Pike MC (1996) Tofu and risk of breast cancer in Asian-Americans. *Cancer Epidemiol Biomark Prev* 5: 901–906

Yamamoto S, Sobue T, Kobayashi M, Sasaki S, Tsugane S (2003) Soy, isoflavones, and breast cancer risk in japan. *J Natl Cancer Inst* 95: 906–913

Yoo K-Y, Tajima K, Kuroishi T, Hirose K, Yoshida M, Miura S, Murai S (1992) Independent protective effect of lactation against breast cancer: a case–control study in Japan. *Am J Epidemiol* 135: 726–733

Yuan JM, Wang QS, Ross RK, Henderson BE, Yu MC (1995) Diet and breast cancer in Shanghai and Tianjin, China. *Br J Cancer* 71: 1353–1358

ELSEVIER

FEMS Immunology and Medical Microbiology 44 (2005) 261–268

FEMS
Immunology and
Medical Microbiology

www.fems-microbiology.org

Identification of *Helicobacter pylori* and the *cagA* genotype in gastric biopsies using highly sensitive real-time PCR as a new diagnostic tool

Shiho Yamazaki [a], Shunji Kato [b], Norio Matsukura [b], Masahiro Ohtani [c], Yoshiyuki Ito [a], Hiroyuki Suto [a], Yukinao Yamazaki [c], Akiyo Yamakawa [a], Shinkan Tokudome [d], Hideaki Higashi [e], Masanori Hatakeyama [e], Takeshi Azuma [a,*]

[a] *Second Department of Internal Medicine, Faculty of Medical Sciences, University of Fukui, Matsuoka-cho, Yoshida-gun, Fukui 910-1193, Japan*
[b] *Department of Surgery for Organ Function and Biological Regulation, Nippon Medical School, Tokyo, Japan*
[c] *Department of Endoscopic Medicine, Faculty of Medical Sciences, University of Fukui, Fukui, Japan*
[d] *Department of Health Promotion and Preventive Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan*
[e] *Division of Molecular Oncology, Institute for Genetic Medicine and Graduate School of Science, Hokkaido University, Sapporo, Japan*

Received 16 September 2004; received in revised form 6 December 2004; accepted 13 December 2004

First published online 7 January 2005

Abstract

The CagA protein is one of the virulence factors of *Helicobacter pylori*, and two major subtypes of CagA have been observed, the Western and East Asian type. CagA is injected from the bacteria into gastric epithelial cells, undergoes tyrosine phosphorylation, and binds to Src homology 2 domain-containing protein-tyrosine phosphatase SHP-2. The East Asian type CagA binds to SHP-2 more strongly than the Western type CagA. Here, we tried to distinguish the CagA type by highly sensitive real-time PCR with the objective of establishing a system to detect *H. pylori* and CagA subtypes from gastric biopsies. We designed primers and probe sets for Western or East Asian-*cagA* at Western-specific or East Asian-specific sequence regions, respectively, and *H. pylori* 16S rRNA. We could detect the *H. pylori* 16S rRNA gene, Western and East Asian-*cagA* gene from DNA of gastric biopsies. The sensitivity and specificity for *H. pylori* infection was 100% in this system. In Thai patients, 87.8% (36/41) were *cagA*-positive; 26.8% (11/41) were Western-*cagA* positive and 53.7% (22/41) were East Asian-*cagA* positive, while 7.3% (3/41) reacted with both types of *cagA*. These results suggest that this real-time PCR system provides a highly sensitive assessment of CagA type as a new diagnostic tool for the pathogenicity of *H. pylori* infection.

Keywords: *Helicobacter pylori*; *cagA*; Real-time PCR

1. Introduction

Helicobacter pylori, a spiral, Gram-negative, microaerophilic bacterium, colonizes at least half of the world's human population and is recognized as a major cause of chronic gastritis, peptic ulcer, and an important risk factor for gastric cancer [1–3]. On the basis of various epidemiological studies, *H. pylori* has been classified as a class I carcinogen in humans by a Working Group of the World Health Organization International Agency for Research on Cancer [4].

CagA protein, which is encoded by the *cagA* gene, is a highly immunogenic protein, and is one of the most studied virulence factors of *H. pylori*. Recent studies

* Corresponding author. Fax: +81 776 61 8110.
E-mail address: azuma@fmsrsa.fukui-med.ac.jp (T. Azuma).

0928-8244/$22.00
doi:10.1016/j.femsim.2004.12.011

have indicated that *H. pylori* strains possessing *cagA* are associated with significantly increased risk for the development of atrophic gastritis and gastric cancer [5–9]. The *cagA* gene is located at one end of a 40-kb DNA insertion known as the *cag* pathogenicity island (*cag* PAI), and may have originated from a non-*Helicobacter* source. The *cag* PAI contains 31 putative genes, 6 of which are thought to encode components of a bacterial type IV secretion system, which specializes in the transfer of a variety of multimolecular complexes across the bacterial membrane to the extracellular space or into other cells [10–12]. Recent studies have provided a molecular basis for the pathological actions of CagA on gastric epithelial cells. After attachment of *cagA*-positive *H. pylori* to gastric epithelial cells, CagA is directly injected from the bacteria into the cells via the bacterial type IV secretion system and undergoes tyrosine phosphorylation in the host cells [13–17]. Tyrosine phosphorylation of CagA occurs at the unique Glu–Pro–Ile–Tyr–Ala (EPIYA) motifs present in the C-terminal region [18–20]. Furthermore, it was recently confirmed that translocated CagA forms a physical complex with Src homology 2 domain-containing protein-tyrosine phosphatase (SHP-2) in a phosphorylation-dependent manner and deregulates its enzymatic activity [20]. SHP-2 is known to play an important positive role in mitogenic signal transduction [21]. In addition, SHP-2 is actively involved in the regulation of spreading, migration, and adhesion of cells [22,23]. Deregulation of SHP-2 by CagA may induce abnormal proliferation and movement of gastric epithelial cells. We have also shown that the CagA–SHP-2 complex is found in in vivo human gastric mucosa [24], suggesting that protein interaction plays a role in the pathogenesis of *cagA*-positive *H. pylori* infection.

H. pylori can be divided into distinct populations with different geographical distributions [12,25]. It has been reported that large sequence differences distinguish the *cagA* gene fragments from Asian strains and other strains [26,27]. The molecular weight of the CagA protein varies between 128 and 140 kDa. Variation in the size of the CagA protein is related to the presence of a variable number of repeat sequences in the 3′ region of the gene [28–31]. The phosphorylation sites are located in the repeat region of CagA [18–20]. Recently, it was also demonstrated that the predominant CagA protein isolated in East Asia, where gastric cancer is prevalent, has a distinct sequence at the region that corresponds to the repeat sequence of Western CagA. After tyrosine phosphorylation, this East Asian-specific sequence confers stronger SHP-2 binding and transforming activities than the Western-specific sequence [27]. Since the potential of CagA to disturb host cell functions as a virulence factor could be determined by the degree of SHP-2-binding activity, the diversity of the CagA phosphorylation site may be an important variable in determining the clinical outcome of infection by different *H. pylori* strains.

Therefore, it may be important to distinguish between East Asian and Western type CagA to expedite clinical procedures, for example, eradication therapy. In the present study, we developed a highly sensitive real-time PCR system as a new diagnostic tool to identify *H. pylori* and the *cagA* gene type in DNA samples from gastric biopsies. Furthermore, we also examined the distribution of CagA diversity in Thai patients.

2. Materials and methods

2.1. Establishment of primers and probes

Primers and probes for *H. pylori* 16S rRNA, *cagA*-Western, and *cagA*-East Asian were designed using "Pri-

Table 1
Sequences of oligonucleotide primers and sequence-specific probe sequences for *H. pylori* 16S rRNA, *cagA*-Western, and *cagA*-East Asian

Gene and oligonucleotide	Sequence	Corresponding DNA sequence
16S rRNA		
Forward primer	5′-TGC GAA GTG GAG CCA ATC TT-3′	1381–1400[a]
Reverse primer	5′-GGA ACG TAT TCA CCG CAA CA-3′	1499–1480[a]
Probe	5′-(FAM) CCT CTC AGT TCG GAT TGT AGG CTG CAA C (TAMRA)-3′	1408–1435[a]
cagA-Western		
Forward primer	5′-AGG CAT GAT AAA GTT GAT GAT-3′	2854–2874[b]
Reverse primer	5′-AAA GGT CCG CCG AGA TCA T-3′	2945–2927[b]
cagA-East Asian		
Forward primer	5′-AAA GGA GTG GGC GGT TTC A-3′	2812–2830[c]
Reverse primer	5′-CCT GCT TGA TTT GCC TCA TCA-3′	2903– 2883[c]
cagA-common probe	5′-(FAM) TCA GCT AGC CCT GAA CCC ATT TAC GCT AC (TAMRA)-3′	2893–2921[b] 2845–2872[c]

[a] Nucleotide positions in the 16S rRNA gene of *H. pylori* 85D08 (GenBank accession no. U00679).
[b] Nucleotide positions in the *cagA* gene of *H. pylori* 26695 (GenBank accession no. AE000569).
[c] Nucleotide positions in the *cagA* gene of *H. pylori* F32 (GenBank accession no. AF202972).

mer Express ver.1.5" (Applied Biosystems, Foster City, CA, USA) (Table 1). The alignment of the deduced amino acid sequence in the 3′ region of the *cagA* gene among strains 26695 and F32, which are typical Western and East Asian-*cagA*, respectively, is shown in Fig. 1. We previously reported that NCTC11637 possesses 5 EPIYA motifs that are potential targets of tyrosine phosphorylation [27]. These EPIYA motifs are involved in the interaction of CagA with SHP-2. The first and second EPIYA motifs (which we designated "EPIYA–A" and "EPIYA–B", respectively) are present in almost all CagA proteins, whereas the remaining 3 EPIYA motifs (which we designated "EPIYA–C") were made by duplication of an EPIYA containing 34-amino acid sequence. Because the 34-amino acid sequence has various numbers, ranging from 1 to 3 in most Western CagA proteins, we designated it the "Western CagA-specific, SHP-2-binding sequence" (WSS) [27]. The WSS contains EPIYA–D1, EPIYA–D2, and EPIYA–D3 motifs, as defined by Covacci et al. [28], or R1 and WSR regions, as defined by Yamaoka et al. [30,31], while 26695 has a single WSS and is thus classified as the "A–B–C" type, whereas 11637-CagA is classified as the "A–B–C–C–C" type. On the other hand, the amino acid sequence of East Asian CagA is quite different from that of Western CagA. The predominant East Asian CagA proteins do not have the WSS, but instead, possess a distinct sequence that we designated "East Asian CagA-specific, SHP-2-binding sequence" (ESS) in the corresponding region [27]. ESS contains a JSR region, which has been previously defined by Yamaoka et al. [30,31], and also possesses an EPIYA motif, designated "EPIYA–D". F32 has a single ESS and is thus classified as the "A–B–D" type (Fig. 1). Therefore, to distinguish between Western and East Asian CagA by PCR, we established primers and probe sets involving the WSS and ESS regions, respectively.

2.2. Subjects

A total of 41 gastric biopsy samples obtained from the greater curvature of the upper gastric body of *H. pylori*-positive patients were used in this study. All patients underwent gastroduodenal endoscopy at Chiang Mai University, Kingdom of Thailand. The patients included 24 with chronic gastritis (13 men and 11 women; mean age, 52.5 years), 4 with gastric ulcer (4 men; mean age, 66.5 years), 5 with duodenal ulcer (4 men and 1 woman; mean age, 52.6 years), and 8 with gastric cancer (6 men and 2 women; mean age, 57.2 years). *H. pylori* infection was diagnosed by Hematoxylin–eosin staining, toluidine blue staining, and *H. pylori*-specific antibody immune staining as described previously [32]. Four gastric biopsy samples obtained from the greater curvature of the upper gastric body of *H. pylori*-negative patients were also used as negative controls in this study. Four controls received eradication therapy more than two years earlier, and were diagnosed as *H. pylori*-negative by histology and urea breath tests at least twice during follow-up.

DNA was extracted from tissues using a 'GeneRelease™' kit (Bioventures Inc., Murfreesboro, TN, USA) and stored at 4 °C until amplification was performed. Although these samples contain almost all host DNA, there are very small concentrations of *H. pylori* DNA if these are present.

2.3. Helicobacter pylori *culture*

In the present study, we used the strains 26695 and F32 as positive controls and standard samples. *H. pylori* isolates were inoculated onto a trypticase soy agar (TSA)-II/5% sheep blood plate and cultured for two to three days at 37 °C under microaerobic conditions (5% O_2, 5% CO_2, 90% N_2). *H. pylori* was harvested from

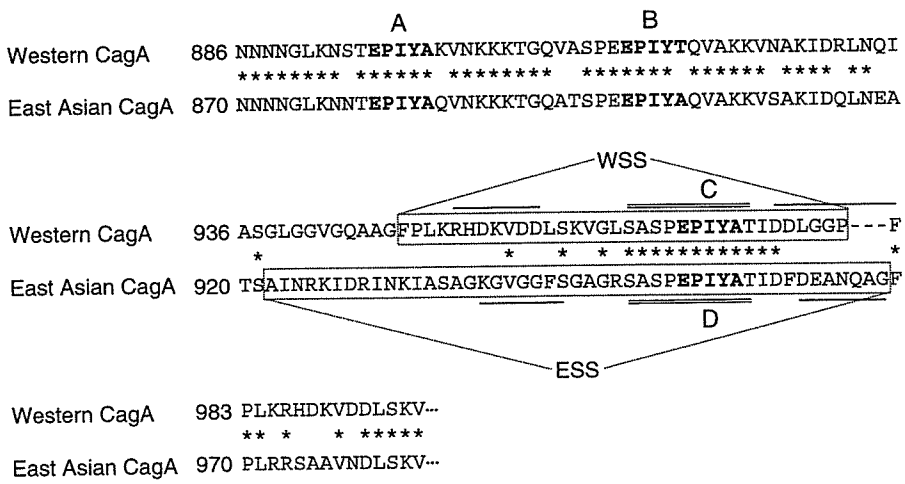

Fig. 1. Alignment of the deduced amino acid sequence of the EPIYA regions for Western (strain 26695) and East Asian CagA (strain F32). Each position of forward and reverse primers is underlined and the probe position is double-underlined.