Cancer registry data based estimation of regional cancer incidence: application to breast and colorectal cancer in French administrative regions

Marc Colonna, Pascale Grosclaude, Jean Faivre, Ali Revzani, Patrick Arveux,
Gilles Chaplain, Brigitte Tretarre, Guy Launoy, Josette Macé Lesec'h, Nicole Raverdy,
Paul Schaffer, Antoine Buémi, François Ménégoz, Roger J Black

Abstract

Study objective—In many countries, cancer registries cover only a small part of the national population. Cancer incidence for the rest of the country has therefore to be estimated. This can be done from mortality data using the relation between incidence and mortality observed in the cancer registry areas. Such an approach was used to study geographical variation and trend of colorectal and breast cancer incidence in France where 10% of the national population is covered by cancer registries.

Design—This study applies the incidence/mortality ratios of cancer registry areas to regional mortality data to obtain an estimation of cancer incidence at a given point in time. Age and period effects are included in the statistical models.

Main results—The incidence estimations are given for 21 administrative regions and three time points (1985, 1990, 1995). The European standardised incidence rates for breast cancer ranged from 86.8 to 128.8. For colorectal cancer, these rates ranged from 48.2 to 79.6 for men, and from 32.5 to 48.8 for women. Breast cancer incidence has increased considerably between 1985 and 1995 with a higher increase in the north than in the south of France. The incidence of colorectal cancer has also increased, albeit to a lesser extent.

Conclusion—The incidence estimation method proposed leads to regional incidence rates that are useful for planning health care services on a regional basis and may also be used to study regional differences in incidence. This method is useful when only partial incidence data are available.

(*J Epidemiol Community Health* 1999;**53**:558–564)

A list of the institutions participating in the study is given at the end of the paper.

Correspondence to:
Dr M Colonna, Registre du Cancer de l'Isère, 21 Chemin des Sources 38240 Meylan, France.

Accepted for publication 29 January 1999

Cancer is responsible for over 130 000 deaths per annum in France, representing about a quarter of the total mortality.[1] Although mortality statistics give an overview of the extent of cancer, they do not provide information on the incidence of the different cancer sites, being unaware of curable cancer types and of other mortality causes of incidents and prevalent cancer cases. Only cancer registries, which record exhaustively incident cases in a given geographical area, can establish reliable morbidity statistics that are essential to health care planning, aetiological or trend studies. As cancer registries do not cover the whole country in many countries, incidence rates have often to be estimated. An estimation method that has been previously applied in France to estimate national incidence[2 3] used mortality data from an area without a cancer registry and the incidence and mortality relation observed in the regional cancer registries. In this paper, this method has been extended to provide regional incidence estimates. Results of breast and colorectal cancer incidence estimation are provided and clearly illustrate the geographical heterogeneity that exists in cancer incidence.

Methods

DATA COLLECTION

The data were obtained from the FRANCIM French cancer registries network, which covers 10% of the French population for breast cancer in women and for colorectal cancer in both sexes. Data were obtained from nine French administrative départements: Bas-Rhin (1975–92), Calvados (1978–92), Côte d'Or (1976–1992 for colorectal cancer, 1982–1992 for breast cancer), Doubs (1978–92), Haut-Rhin (1988–1992), Herault (1986–92), Isère (1979–92), Somme (1982–92), and Tarn (1982–1992).

The data on mortality in each region, year of death, age group, sex and site of the tumour were available for the 1975–1992 period. They are provided by the Institut National de la Santé et de Recherche Médicale (INSERM). As practically no incident cases are observed before the 20–24 years age group for both breast and colorectal cancer, the age range starts at 20 years.

The population data estimated with respect to age, year, sex, and region were supplied by the Institut National de la Statistique et des Etudes Economiques (INSEE). Table 1 gives the breakdown of the population by sex and region for 1992. In table 1, the percentage of persons aged 75 years and over is also mentioned. It can be seen that there was considerable diversity in the age structures of the regional populations.

STUDY DESIGN

Modelisation

The previous studies[2 3 5-7] applied the incidence/mortality ratios of cancer registry areas to national (for example, France) mortality data to obtain an estimation of cancer incidence at a given point in time. In a recent

Table 1 Breakdown of total population by sex and regions in 1992. Percentages of men and women over 75 years old

1992	Whole population (men)	Whole population (women)	% 75+ (men)	% 75+ (women)
Alsace	805 028	844 141	3.4	7.3
Aquitaine	1 366 365	1 461 691	5.7	9.9
Auvergne	641 997	676 429	5.6	10.1
Basse Normandie	681 460	717 718	4.1	8.2
Bourgogne	789 167	826 975	5.7	10.0
Bretagne	1 366 641	1 446 451	4.5	8.9
Centre	1 174 379	1 224 403	5.6	9.3
Champagne	662 033	686 686	4.3	8.0
Franche Comté	544 828	558 154	4.4	8.0
Haute Normandie	857 736	898 086	3.8	7.3
Languedoc R	1 045 453	1 113 210	6.1	9.7
Limousin	349 105	372 456	7.6	12.7
Lorraine	1 130 878	1 175 224	3.7	7.3
Midi Pyrénées	1 201 232	1 260 054	6.2	10.1
Nord	1 929 205	2 045 053	3.2	6.8
Pays de Loire	1 506 920	1 580 311	4.5	8.3
Picardie	902 701	927 298	3.7	7.0
Poitou Charente	784 460	820 529	6.3	10.1
PACA	2 088 994	2 246 857	5.6	9.5
Ile de France	5 242 685	5 561 017	3.4	6.8
Rhône Alpes	2 668 829	2 782 482	4.0	7.7

study, the mortality trend, the trends in mortality and in incidence/mortality ratios in the total registered population were used to estimate national incidence rates over a 20 year period and therefore an estimate of incidence trend. This method has been extended to regional incidence. The procedure was divided into four main stages:

(1) For each combinations of sex and site, we fitted a log-linear model for the age (subscript *i*), population (subscript *j*) and period (subscript *t*) specific mortality rates λ_{ijt}, in the national and pooled registry populations expressed by the following relation:

$$\log(\lambda_{ijt}) = a_i + p_j + \beta_i t + \gamma_i t^2$$
$$i=1,...,14$$
$$j=1, 2 \tag{1}$$

where a_i, β_i, γ_i and p_j are categorial variables coding for age groups (20–24, 25–29,... 80–84 and 85+) and populations (national and registry), *t* is coding for the single years 1975–92. The model was fitted to obtain smoothed values for national and pooled registries mortalities. The model allows rates to differ in both populations,

subject to the constraint of a common rate ratio for all age groups and identical age specific time trends. The effect of the larger national population is to smooth the sparse age specific data of the registry populations. The inclusion of second degree polynomial term allows curvature in the age specific time trends. This seemed to be adequate for representing the slight and slowly changing trends observed in cancer mortality in large populations.

(2) For each combination of sex and site, we fitted log-linear models for the age (subscript *i*), population (subscript *j*) and period (subscript *t*) specific mortality rates λ_{ijt} in the 21 French areas, expressed by the relation:

$$\log(\lambda_{ijt}) = a_i + \beta_i t + \gamma_i t^2 + p_j + q_j t + r_j t^2$$
$$i=1..14$$
$$j=1..21 \tag{2}$$

where the parameters have the same meaning as in model (1) except that *j* codes now for the 21 different French areas. The model was fitted to obtain smoothed values for the local mortalities on the same assumptions as in model (1) with a possible area specific time trend.

(3) The observed incidence and mortality data from the pooled registries seemed unsuitable for estimating age and year specific incidence/mortality ratios as these observations may be sparse for some age groups. Models were therefore fitted to estimate trends in the age specific incidence/mortality ratios in the pooled registry populations, expressed by the relation:

$$\log(c_{it}) = \log(\hat{d}_{it}) + a_i + \beta_i t \tag{3}$$

where c_{it} is the number of observed incident cases in the pooled registry populations at age *i* and year *t* and $\hat{d}_{it}$ is the estimated deaths in these populations from model (1). The model assumes a log-linear relation between incidence and mortality. A quadratic term may be included for some particular sites when sudden changes in diagnostic or treatment techniques appeared, as it has been shown for prostatic carcinoma.

(4) Finally, for each cancer site, local incidence estimates were obtained by applying the estimated incidence/mortality ratios from model (3):

$$\frac{\hat{c}_{it}}{\hat{d}_{it}} \tag{4}$$

to the fitted values for local mortality from model (2). Fitted values for the age and year specific local area mortality are required instead of the observed values to avoid extreme variability because of the small number of observed deaths in age groups for some sites. Estimated number of incident cases $\hat{c}_{ijt}$ in the age group *i*, area *j* and year *t* are:

$$\hat{c}_{ijt} = \hat{d}_{ijt} \cdot \frac{\hat{c}_{it}}{\hat{d}_t} \tag{5}$$

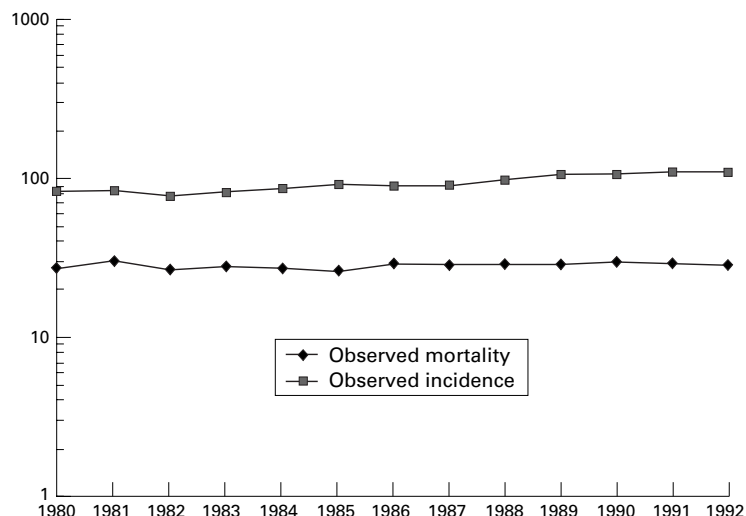where $\hat{d}_{ijt}$ is the number of estimated mortality cases from model (2).

Figure 1 Plot of observed incidence and mortality European age standardised ratio against time for the French départements covered by a cancer registry—breast cancer.
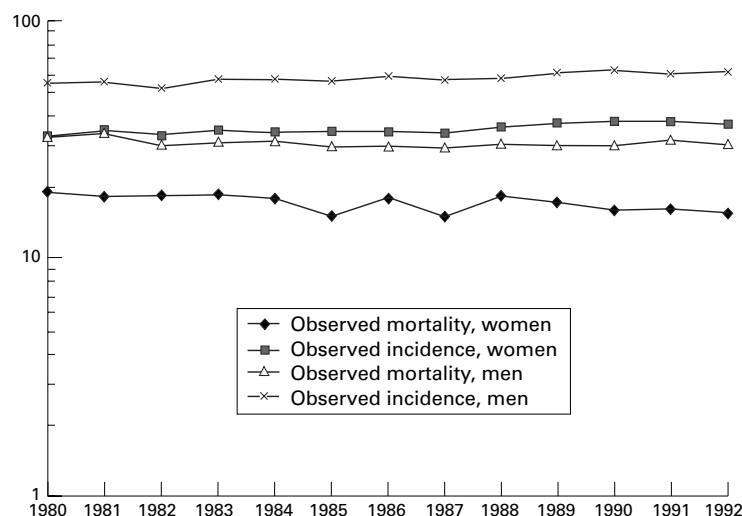
Figure 2   Plot of observed incidence and mortality European age standardised ratio against time for the French départements covered by a cancer registry—colorectal cancer for women.

Assessment of the reliability of the method
To check the reliability of the estimations, we have used some form of cross validation. We have dropped each cancer registry data from the analysis in turn and estimated its incidence from the others. Then we have compared observed and estimated incidence from each of the nine modelisations we have performed.

Variance estimation of the estimates
To estimate the variance of the estimated number of cases, we used the formulation (5). We have computed the variance of $\hat{c}_{ijt}$ using a combination of the variance of the predicted values $\hat{d}_{ijt}$ from model (2) and of the variance of the ratio $\frac{\hat{c}_{it}}{\hat{d}_{it}}$. The variance of $\hat{d}_{ijt}$ was obtained using the classic variance of predicted value;

the variance of $\frac{\hat{c}_{it}}{\hat{d}_{it}}$ is the empirical variance of the estimated values of this ratio obtained for the registry in turn in the reliability assessment. The combination takes the following form:

$$\mathrm{var}(\hat{c}_{ijt}) = (\frac{\hat{c}_{it}}{\hat{d}_{it}})^2 \mathrm{var}(\hat{d}_{ijt}) + \hat{d}_{ijt}^2 . \mathrm{var}(\frac{\hat{c}_{it}}{\hat{d}_{it}}) \quad (6)$$

$$\mathrm{var}(\hat{c}_{\cdot jt}) = \sum_i \mathrm{var}(\hat{c}_{ijt}) \quad (7)$$

The GLIM statistical package[8] was used for calculations and to set up the models. European age standardised incidence rates were computed to take into account the differences in age structure between French areas.

Results
INCIDENCE/MORTALITY
Our estimations of incidence are based essentially on the incidence/mortality ratios. Figures 1 and 2 show the plots of the components of these ratios over the period 1980–1992 for the two cancer sites and for the nine French départements covered by a cancer registry. For each cancer site, the comparison between age standardised mortality rates and age standardised incidence rates shows a definite increase in the incidence/mortality ratio over this 12 year period. The most distinct increase was for breast cancer. The underlying hypothesis of a log-linear trend of the observed incidence/mortality thus appeared to be verified for the two studied sites.

REGIONAL INCIDENCE OF BREAST AND
COLORECTAL CANCER IN 1992
Tables 2, 3, and 4 list the estimated number of cases of breast cancer in women and of colorectal cancer in both men and women for 1992 for each age group in each region. The standard deviation of estimated number of cases for each

Table 2   Number of cases of breast cancer in women in the different age groups with European age standardised incidence and mortality rates in 1992. Number of cases and European standardised incidence rates in 1985, 1990, and 1995

	1992								1985		1990		1995	
	20–44*	45–64*	65–74*	75–84*	85+*	all ages*	EASIR	EASMR	all ages*	EASIR	all ages*	EASIR	all ages*	EASIR
Alsace	144	411	210	118	38	921 (13.44)	107.7	27.3	760	95.1	875	104.8	1009	111.4
Aquitaine	229	721	427	249	94	1720 (20.41)	101.2	27.5	1244	77.9	1557	93.3	2059	115.7
Auvergne	109	356	209	123	45	842 (12.46)	105.5	29.9	635	81.8	777	98.6	978	117.8
Basse Normandie	112	345	198	107	37	799 (12.21)	104.2	29.8	595	82.5	731	97.4	948	117.4
Bourgogne	122	383	228	136	54	923 (13.02)	97.9	26.8	773	85.6	888	95.8	974	99.1
Bretagne	195	622	379	209	63	1468 (18.47)	90.5	26.5	1141	74.8	1371	86.2	1652	96.5
Centre	185	547	309	188	72	1301 (16.57)	97.1	26.4	1057	84.6	1230	94.0	1420	100.8
Champagne	106	298	159	92	34	689 (11.08)	97.7	26.8	532	78.9	635	91.8	794	107.3
Franche Comté	78	237	125	71	24	535 (9.51)	91.5	26.5	453	82.6	515	90.2	579	93.8
Haute Normandie	159	433	224	125	45	986 (14.15)	110.8	32.7	783	95.2	925	106.6	1093	116.8
Languedoc Roussillon	155	513	313	172	62	1215 (15.95)	92.9	25.6	1035	85.9	1175	92.2	1273	92.2
Limousin	52	182	121	77	30	462 (8.43)	95.6	25.7	345	71.6	418	87.3	549	110.5
Lorraine	193	573	294	160	52	1272 (17.09)	105.0	28.4	981	85.3	1176	99.0	1476	115.8
Midi Pyrénées	167	542	315	189	70	1283 (16.36)	86.8	21.8	1064	75.9	1224	84.3	1384	89.7
Nord	405	1131	609	321	102	2568 (28.74)	128.8	37.2	1979	103.6	2362	120.5	3008	143.8
Pays de Loire	246	718	396	226	83	1669 (20.23)	100.8	29.2	1314	85.8	1553	96.4	1902	108.0
Picardie	166	448	231	125	45	1015 (14.69)	112.0	29.7	759	89.9	929	105.1	1191	124.7
Poitou Charente	118	377	224	133	50	902 (12.93)	94.5	28.8	708	78.8	846	90.5	1000	99.4
PACA	337	1098	630	363	125	2553 (27.89)	97.8	25.1	1979	82.5	2386	93.7	2839	103.8
Ile de France	1093	2814	1199	718	292	6116 (61.00)	114.0	32.2	4832	95.4	5732	108.7	6807	122.4
Rhône Alpes	457	1322	657	376	139	2951 (31.75)	102.9	27.9	2385	91.3	2802	100.6	3186	105.3
Total	4828	14071	7457	4278	1556	32190	103.7	28.7	25354	87.1	30107	99.1	36121	111.0

*Estimated number of new cases. EASIR : European Age Standardised Incidence Rate. EASMR : European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases.

Table 3   Number of cases of colorectal cancer in women in the different age groups with European standardised incidence and mortality rates in 1992. Number of cases and European standardised incidence rates in 1985, 1990, and 1995

	1992								1985		1990		1995	
	20–44*	45–64*	65–74*	75–84*	85+*	all ages*	EASIR	EASMR	all ages*	EASIR	all ages*	EASIR	all ages*	EASIR
Alsace	15	104	125	129	57	430 (8.99)	40.7	16.4	378	39.0	415	40.4	455	40.6
Aquitaine	24	177	246	262	135	844 (14.78)	36.7	15.7	656	31.3	767	34.3	1024	42.4
Auvergne	9	81	110	119	59	378 (8.13)	35.0	17.4	331	32.4	361	34.0	421	37.1
Basse Normandie	9	76	103	102	48	338 (7.71)	33.8	16.0	288	31.8	320	33.1	390	36.7
Bourgogne	12	101	140	152	82	487 (9.81)	37.7	17.2	365	30.3	434	34.4	612	45.3
Bretagne	23	179	249	250	104	805 (14.59)	37.6	15.5	659	34.1	753	36.4	920	40.3
Centre	20	142	189	209	110	670 (12.79)	37.3	16.9	528	32.4	615	35.5	795	42.0
Champagne	10	73	92	98	49	322 (7.39)	35.6	16.7	282	33.5	306	34.8	357	37.5
Franche Comté	8	56	70	73	34	241 (5.93)	32.5	12.9	226	33.3	237	33.0	251	31.6
Haute Normandie	16	111	137	139	69	472 (9.96)	42.6	17.6	333	33.4	417	38.9	604	51.7
Languedoc Roussillon	15	127	181	182	90	595 (11.54)	33.8	16.1	455	28.6	547	32.1	694	37.2
Limousin	4	47	70	83	43	247 (6.08)	35.0	16.7	208	30.4	227	32.7	297	40.3
Lorraine	22	158	188	187	83	638 (12.55)	42.7	19.7	472	34.5	565	39.0	816	51.3
Midi Pyrénées	17	140	189	207	105	658 (12.13)	32.9	15.1	533	29.0	605	31.0	776	36.8
Nord	44	292	363	351	152	1202 (21.32)	48.8	20.5	963	41.9	1106	46.1	1440	55.2
Pays de Loire	25	171	220	229	115	760 (13.76)	35.5	15.8	636	33.1	720	34.9	841	36.6
Picardie	16	103	126	125	61	431 (9.38)	38.6	17.0	349	33.6	399	36.7	504	42.9
Poitou Charente	12	104	144	156	80	496 (9.93)	37.9	17.9	383	32.2	450	35.6	604	43.6
PACA	32	251	342	356	168	1149 (19.78)	33.4	15.4	889	29.0	1058	31.8	1331	36.6
Ile de France	108	593	622	677	375	2375 (41.92)	37.5	16.6	2072	34.9	2269	36.5	2594	39.4
Rhône Alpes	47	308	367	383	193	1298 (22.47)	36.4	15.7	1067	33.6	1218	35.3	1460	38.6
Total	488	3394	4273	4469	2212	14836	37.3	16.5	12073	33.2	13789	35.7	17186	41.0

*Estimated number of new cases. EASIR : European Age Standardised Incidence Rate. EASMR : European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases.

area is given. The three tables also show the standardised incidence and mortality rates derived from the European standard population.

There was some disparity in incidence between regions for breast cancer. An incidence above national average was noted in the Nord département, while Bretagne and Midi-Pyrenees had below average incidences. Overall, the estimated number of new cases in 1992 was a slightly over 32 000. The proportion of cases diagnosed before 45 years of age is 15%, 44% between 45 and 64 years, 23% between 65 and 74 years, 13% between 75 and 84 years, and 5% after 84 years. These proportions varied little from region to region. The European standardised mortality rates for breast cancer were almost four times lower than the European standardised incidence rates.

There was less regional disparity in incidence of colorectal cancer for women than for men. In this lastest group, the incidence rates were 10% above national average in Alsace, Nord, Lorraine and Bretagne, but 10% under in Provence Alpes Côte d'Azur, Midi-Pyrenees, Languedoc Roussillon and Franche Comté.

The incidence of colorectal cancer was slightly higher in men than in women (16 401 estimated cases versus 14 836 in 1992). The sex ratio calculated from the incidence rates ranged from 1.3 to 2 depending on the region. Colorectal cancer was sparse before 45 years with 3% of cases in both men and women. The proportion of cases diagnosed at 85 years and over was 7% in men and 15% the women.

TREND OF REGIONAL INCIDENCE OF COLORECTAL AND BREAST CANCER

In an attempt to evaluate the trends in cancer incidence, we have indicated in tables 2, 3, and 4 the number of estimated cases and

Table 4   Number of cases of colorectal cancer in men in the different age groups European standardised incidence and mortality rates in 1992. Number of cases and European standardised incidence rates in 1985, 1990, and 1995

	1992								1985		1990		1995	
	20–44*	45–64*	65–74*	75–84*	85+*	all ages*	EASIR	EASMR	all ages*	EASIR	all ages*	EASIR	all ages*	EASIR
Alsace	22	180	186	124	32	544 (11.30)	79.6	38.1	422	68.3	496	75.6	654	88.4
Aquitaine	27	248	376	261	73	985 (16.15)	59.9	27.7	780	52.4	902	56.8	1164	66.4
Auvergne	12	121	177	123	35	468 (9.42)	60.8	30.7	380	52.1	429	57.2	560	68.5
Basse Normandie	12	116	168	100	24	420 (9.18)	61.0	29.9	299	49.2	368	56.0	551	73.6
Bourgogne	15	146	213	156	47	577 (10.72)	62.2	30.8	469	58.8	531	58.8	682	69.1
Bretagne	28	259	374	237	51	949 (16.27)	65.2	30.5	732	56.5	858	61.4	1157	72.9
Centre	23	202	290	218	67	800 (14.10)	60.0	29.0	635	52.2	736	57.2	927	65.0
Champagne	12	110	142	100	28	392 (8.47)	61.4	31.6	318	53.0	362	58.5	462	67.8
Franche Comté	8	79	102	70	19	278 (6.45)	51.7	27.3	253	52.1	267	51.7	304	52.6
Haute Normandie	15	122	157	101	26	421 (8.94)	54.7	26.2	354	51.4	396	53.7	461	55.8
Languedoc Roussillon	16	167	270	186	53	692 (12.07)	52.2	25.1	557	47.2	642	50.6	791	55.5
Limousin	4	64	105	81	25	279 (6.43)	54.9	29.9	262	52.7	272	54.4	296	55.4
Lorraine	23	213	248	156	38	678 (13.40)	65.4	32.3	538	57.2	614	61.6	824	73.7
Midi Pyrénées	20	191	289	217	64	781 (13.35)	51.9	25.8	653	47.1	726	49.8	911	57.0
Nord	44	341	428	244	62	1119 (20.11)	69.4	32.8	910	65.9	1028	65.9	1327	77.2
Pays de Loire	29	250	348	237	63	927 (16.44)	61.0	31.1	716	53.7	851	58.7	1086	65.1
Picardie	18	143	184	115	33	493 (10.29)	61.1	29.9	397	53.2	455	58.5	581	67.1
Poitou Charente	14	147	227	170	50	608 (11.05)	61.1	34.0	492	54.5	564	58.9	690	64.3
PACA	32	307	451	321	87	1198 (19.33)	48.2	24.4	981	44.7	1129	47.4	1307	49.0
Ile de France	108	755	752	529	175	2319 (41.66)	55.4	29.0	1966	51.0	2174	53.6	2608	59.3
Rhône Alpes	53	432	536	354	98	1473 (24.54)	58.7	28.0	1198	54.4	1373	57.2	1673	61.7
Total	535	4593	6023	4100	1150	16401	58.6	29.0	13312	52.4	15173	56.2	19016	63.5

*Estimated number of new cases. EASIR : European Age Standardised Incidence Rate. EASMR : European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases.

Table 5   Number of observed cases with confidence intervals in each cancer registry in 1992, breast cancer in women, number of estimated cases from cross validation method, and number of observed cases and estimated cases ratios

Registry number	Number of observed cases	Confidence intervals of observed number of cases	Number of estimated cases	observed/ estimated
1	546	502, 594	539	1.01
2	351	316, 390	325	1.08
3	252	223, 285	232	1.09
4	214	187, 245	234	0.91
5	384	347, 424	377	1.02
6	588	542, 638	495	1.19
7	602	556, 652	582	1.03
8	249	220, 282	317	0.79
9	197	171, 227	179	1.10

Table 6   Number of observed cases with confidence intervals in each cancer registry in 1992, colorectal cancer in women, number of estimated cases from cross validation method, and number of observed cases and estimated cases ratios

Registry number	Number of observed cases	Confidence intervals of observed number of cases	Number of estimated cases	observed/ estimated
1	240	211, 272	241	1.00
2	153	131, 179	143	1.07
3	108	89, 130	110	0.98
4	99	81, 121	91	1.09
5	192	167, 221	179	1.07
6	184	159, 213	202	0.91
7	255	226, 288	232	1.10
8	106	88, 128	135	0.79
9	87	71, 107	94	0.93

Table 7   Number of observed cases with confidence intervals in each cancer registry in 1992, colorectal cancer in men, number of estimated cases from cross validation method, and number of observed cases and estimated cases ratios

Registry number	Number of observed cases	Confidence intervals of observed number of cases	Number of estimated cases	observed/ estimated
1	304	272, 340	294	1.03
2	168	144, 195	155	1.08
3	150	128, 176	148	1.01
4	88	71, 108	98	0.90
5	227	199, 259	238	0.95
6	255	226, 288	235	1.08
7	261	231, 295	250	1.04
8	131	110, 155	158	0.83
9	127	107, 151	128	0.99

KEY POINTS
- Incidence/mortality ratio, taking into account age and period efects, is used to estimate regional cancer incidence.
- The method is useful for countries lacking nationwide coverage by cancer registries.
- There are some regional variations in the incidence of breast and colorectal cancer.
- There is an increase in incidence for both cancer sites, with regional differences.

standardised incidence rates based on the European populations for 1985, 1990, and 1995.

In 10 years, the incidence of breast cancer has increased considerably but with regional differences. Increases of more than a third were noted in Aquitaine, Auvergne, Basse Normandie, Champagne, Limousin, Lorraine, Nord and Picardie. Smaller increases were observed in the other regions.

The increases in incidence of colorectal cancer in both men and women have been relatively modest, with a mean increase in incidence of 20%. Incidence remained stable over this decade in: Franche Comté, Haute Normandie, Limousin and Provence Alpes Côte d'Azur for men, and in Alsace, Franche Comté and Pays de Loire for women.

RELIABILITY ASSESSMENT
In tables 5, 6, and 7, the number of observed and estimated cases in each cancer registry are presented. Each cancer registry estimations are computed without taking into account its own observations. The confidence intervals of observed cases are provided. The ratio number of observed cases/number of estimated cases are computed. Except for cancer registry number 8 for the three cancer sites and for cancer registry number 6 for breast cancer,

these ratio lie between 0.9 and 1.1. Apart from those two previous situations, the number of estimated cases are inside the corresponding confidence intervals computed according to the number of observed cases.

Discussion

We provide in this study the first ever estimation of the incidence of cancers in different regions in France. The incidences of breast and colorectal cancer were used as typical examples as they are common sites.

The quality of the estimation of incidence depends on the fit of the models, which in turn relies on the underlying hypotheses. A feature of our methodology was the procedure for smoothing mortality data in the départements covered by a cancer registry. With the assumption of a parallelism in the trend of adjusted mortality rates (model (1)), the data in pooled registries areas were considered as being representative of national data. This actually may not be the case because of geographical differences in the catchment areas of the regional cancer registries. This can be resolved by modelling only the mortality data of the départements covered by a regional cancer registry in the model (1), although it may not be valid when mortality is relatively low. Using this latter method for breast and colorectal cancer, we obtained similar estimations. Model (2) contained parameters for modelling the spatio-temporal trends in regional mortality. Model (3) could be improved by the inclusion of a second degree term in the time variable. This correction is not statistically useful for the cancer sites studied here, although it could be used for other sites, such as the prostate, which has presented rapid changes in incidence over the past decades. As each of the estimated models leads to a good statistical fit, we considered that this methodology is one possible approach for the estimation of regional incidence rates.

Our reliability assessments confirm that the method leads to reasonably good results. An internal study of French cancer registry seems to conclude that the discrepancy showed for cancer registry number 8 is attributable to a cancer registration problem. For breast cancer in cancer registry number 6, we found that incidence level is the highest of the nine French cancer registries and cancer mortality is at the same level as in other cancer registries so that no clear reason can be given to explain the discrepancy between observation and estimation in that area.

The data presented in table 1 show that the estimations involved regions with considerable differences in population size. This means that the variability in the estimations is not constant from region to region, which may lead to a bias in classifying the least populated regions. Further analysis using Bayesian methods would perhaps aid comparison between regions differing in population size. It may be useful too to give confidence intervals of estimated numbers of cases and standardised rate estimates. These estimates were obtained from successive models so that the analytic expression of their confidence intervals are not readily defined. To evaluate the variability of our estimations, we have provided a variance estimation of the total number of estimated cases in each areas.

The use of incidence/mortality ratios to estimate regional incidence is based on the assumption of comparable survival rates in the different regions. This assumption is largely validated by the data from the EUROCARE study. The survival rates in the French départements that participated in this study were comparable and have provided the overall data for France.[9] One should bear in mind, however, that targeted public health measures such as screening may affect the validity of this assumption in the future. The fact that the estimates of regional incidence rates for cancer rely on certain simplifying assumptions means that the values obtained should be viewed as orders of magnitude. The extrapolations to 1995 for the regions considered here produced values that were slightly higher than those of a previous national evaluation.[4] This was attributed to the fact that the data in this study were obtained over a longer period of time, especially for the Bas-Rhin département. The extrapolations should thus be interpreted with due caution and may need regular updating.

Estimations of incidence are crucial for health care planning. Knowledge of the expected number of cases in each region, the available data on care facilities, taking the age factor into account, can provide an estimate of needs for surgical beds or of the number of people likely to require chemotherapy and radiotherapy. Estimations of regional incidence produce objective data on which health care needs in oncology may be relied on.

The estimations of incidence also highlight the limitations of mortality statistics for determining the magnitude of the problems represented by the two types of cancer studied here. For breast cancer, the incidence rates were four times higher than mortality rates, and even for colorectal cancer, incidence rates were double those of mortality. The statistics of incidence and mortality were also complementary. The increase in the incidence/mortality ratio is indicative of an improved prognosis for the two types of cancer studied. The increase was more pronounced for breast than for colorectal cancer. This trend has also been reported recently in Sweden for breast cancer[10] and in the USA for colorectal cancer.[11] The improved prognosis may stem from an earlier diagnosis or the advent of more effective treatments, or both.

Further analysis of registry data could provide more information on this point.

The results of this study indicate that there were some regional variations in the incidence of breast cancer in France, the incidence ratio of the highest risk region to the lowest is only 1.5. In general, the incidence of breast cancer is lower than average in the south and above average in the north. A same ratio is estimated for colorectal cancer for women. For men, we found a ratio of highest to lowest risk incidences of 1.7. Incidence was particularly high in Alsace, and above average in Lorraine and Nord regions. The lowest incidences were found in Provence Alpes Côte d'Azur. The role of the Mediterranean diet has been proposed to account for the lower incidence in the south than in the north of France. A diet rich in fruit and vegetables has been shown to reduce the risks of both types of cancer studied.

The analysis of time trends is expected to provide valuable information, although more detailed analysis will require data collected over a longer period of time. Already it seems that there are regional differences in such trends. The changes in incidence of breast cancer were found to be more pronounced in the north than in the south of France. The incidence of colorectal cancer has also increased but to a smaller extent than for breast cancer. In certain regions the incidence has remained stable.

Our study shows that there is a heterogeneity both in spatial distribution and in time trends of breast and colorectal cancer incidence. The results of this study are therefore helpful to plan regional health care needs when no national cancer incidence registration exists. The method we propose uses the empirical relation between incidence and mortality for each period and age class. This method is useful for countries lacking nationwide coverage by cancer registries when no great heterogeneity exists between areas in survival. It is an interesting alternative method to that used by Verdecchia *et al*[12] when no reliable survival data are available. Of course, our method may be improved. Firstly, our method uses observed correlation between incidence and mortality in areas covered by cancer registries at each time point where observations are avalailable so that no time lag between incidence and mortality is taken into account; it may be interesting to evaluate the effect of including such a lag. Secondly, our mortality smoothing does not include any spatial dependence; perhaps such information may improve smoothed mortality rates. Lastly, no mortality variance is included in our model (3), where each estimated mortality case is an offset.

1 Lion J, Hatton F, Maguin P, *et al. Statistiques des causes médicales de décès, 1990.* Paris: Editions INSERM, 1993.
2 Benhamou E, Laplanche A, Wartelle M, *et al.* Incidence des cancers en France, 1978–1982. Paris: Editions INSERM, 1990.
3 De Vathaire F, Koscielny S, Rezvani A, *et al.* Estimation de l'incidence des cancers en France, 1983–87. Paris: Editions INSERM, 1996.
4 Ménégoz F, Black R, Arveux P, *et al.* Cancer incidence and mortality in France in 1975–95. *Eur J Cancer Prev* 1997;**6**: 442–66.
5 Moller Jensen O, Estève J, Moller H, *et al.* Cancer in the European Community and its Member States. *Eur J Cancer* 1990;**26**:1167–256.
6 Parkin D, Pisani P, Ferlay J. Estimates of the worldwide incidence of eighteen major cancers in 1985. *Int J Cancer* 1993;**54**:594–606.
7 Estève J, Kricker A, Ferlay J, *et al. Facts and figures of cancer in the European Community.* Lyon: International Agency for Research on Cancer, 1993.
8 Francis B, Green M, Payne C. *The Glim system. Release 4 manual.* Oxford: Clarendon Press, 1993.
9 Berrino F, Sant M, Verdecchia A, *et al. Survival of cancer patients in Europe, the Eurocare Study.* Lyon: IARC Scientific Publications n132, 1995.
10 Garne J, Aspergren K, Balldin G, *et al.* Increasing incidence of declining mortality from breast carcinoma. *Cancer* 1997; **79**:69–74.
11 Chu C, Tarone R, Chow W, *et al.* Temporal patterns in colorectal cancer incidence, survival and mortality from 1950 through 1990. *J Natl Cancer Inst* 1994;**86**:997–1006.
12 Verdecchia A, Capocaccia R, Egidi V, *et al.* A method for the estimation of chronic disease morbidity and trends from mortality data. *Stat Med* 1989;**8**:201–16.