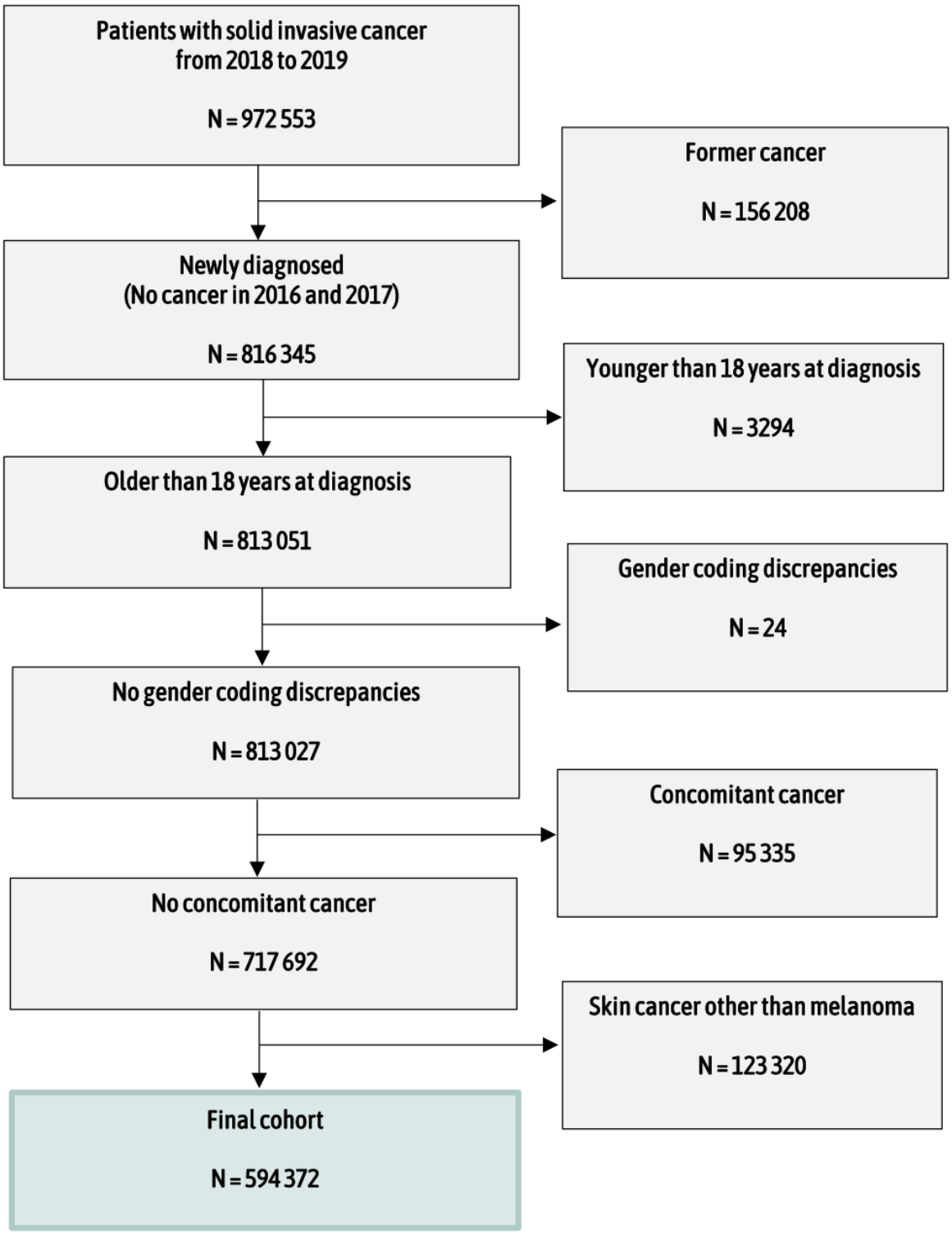
Supplementary material

Supplementary material 1. Flow-chart for patient selection



Supplementary material 2. STROBE Statement

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5-6
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	5-6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	8-9
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	9
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	5-6

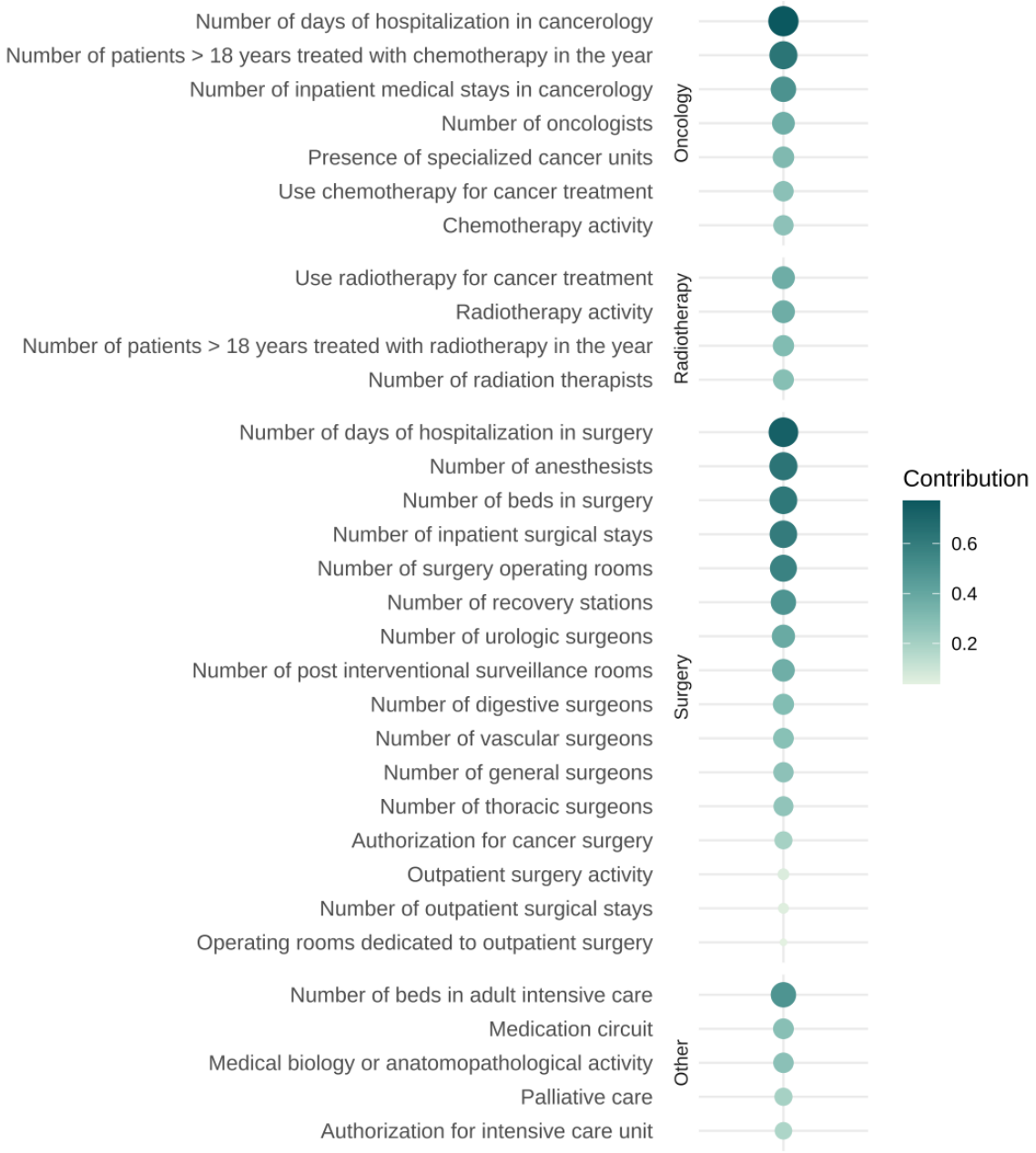
		(c) Consider use of a flow diagram	5-6
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10-11
		(b) Indicate number of participants with missing data for each variable of interest	21
		(c) Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	Report numbers of outcome events or summary measures over time	10-13
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-13
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	13
Discussion			
Key results	18	Summarise key results with reference to study objectives	13-14
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-14
Generalisability	21	Discuss the generalisability (external validity) of the study results	15-16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	9

Supplementary material 3. Details on the methodology of comorbidities

We relied on previously published articles[1–10] to define the list of considered comorbidities, together with the medical codes used to identify them in the French national health insurance database. Diagnosis codes were recorded with the International Classification of Diseases—10th revision, ICD-10[11]. Procedures were recorded with the CCAM classification ("Classification Communes des Actes Médicaux"). The final list of comorbidities included 51 pathologies, gathered into 12 categories: (1) Cardiovascular, (2) Endocrine, (3) Frailty, (4) Gastrointestinal, (5) Immune, (6) Kidney, (7) Liver, (8) Neurologic, (9) Psychiatric disorders, (10) Pulmonary, (11) Rheumatologic disease and connectivopathies, and (12) Other. A patient was suspected to suffer from a given comorbidity at the time of cancer diagnosis if there was at least one CCAM procedure code or ICD-10 diagnosis code associated with the given comorbidity in the year preceding the date of cancer diagnosis.

Supplementary material 4. List of the 32 variables and their contribution to the oncological and surgical expertise score

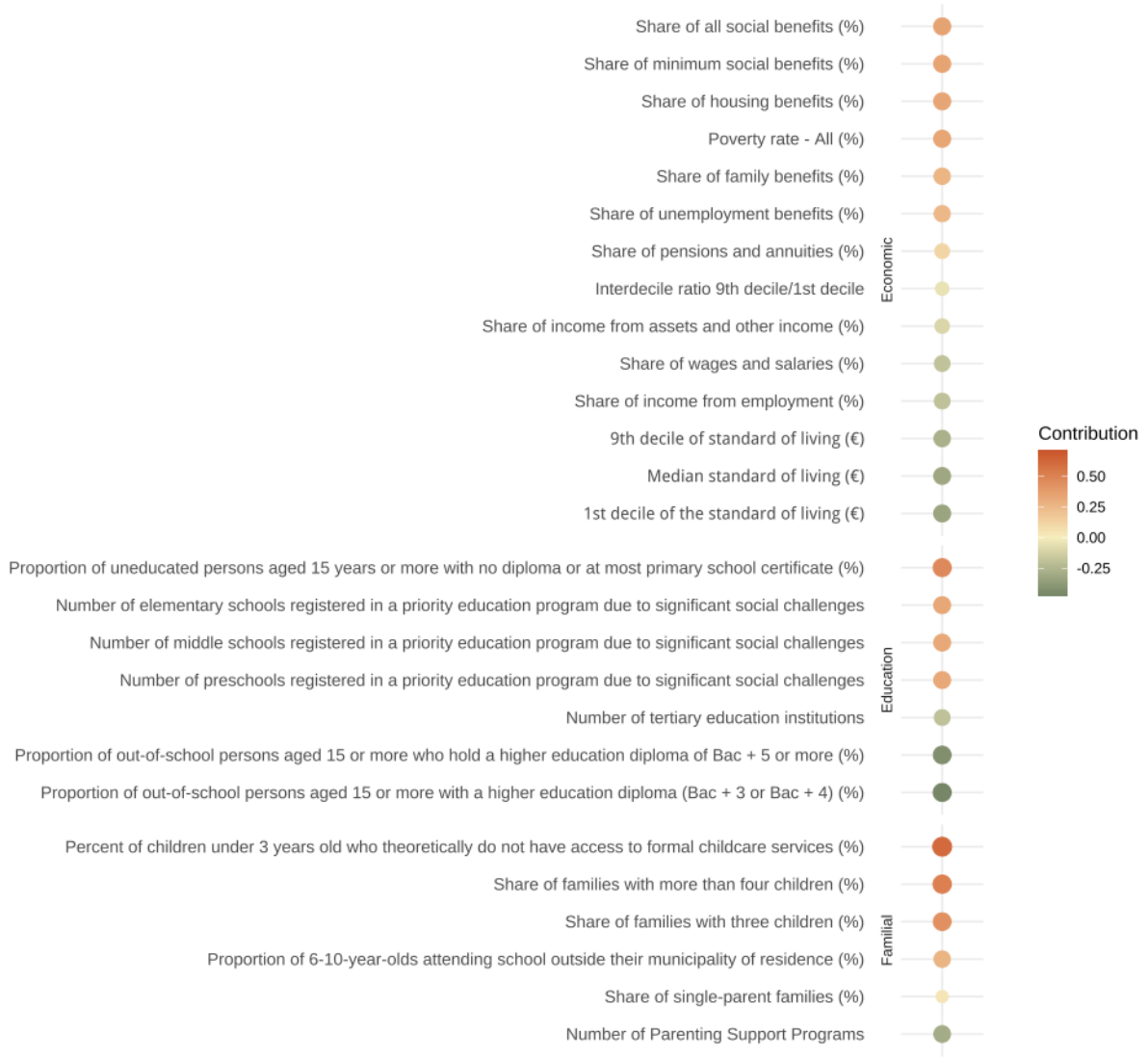
We utilized Factor Analysis of Mixed Data (FAMD) to calculate an expertise score for each centre, which ranged from 0 (low expertise) to 1 (high expertise). The variables were scaled prior to computing the score, and the first Principal Component was used for this calculation. To improve the normality of the data distribution, we applied a logarithmic transformation to the computed score.

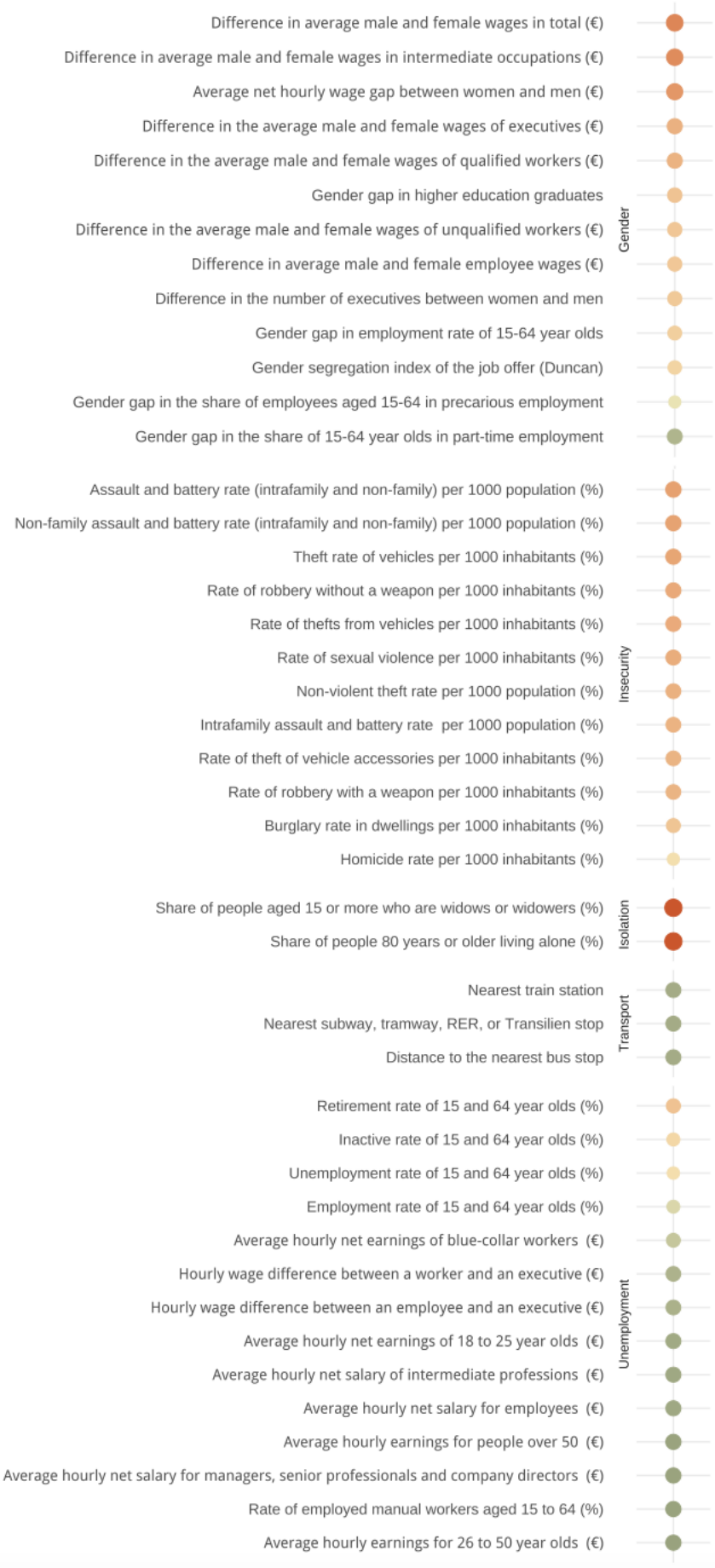


Supplementary material 5. List of the 71 socio-environmental variables and their contribution to the 8 socio-environmental indices (Economic deprivation, Education barriers, Familial hardship, Gender-related wage disparities, Insecurity, Social Isolation, Inaccessibility to Public Transportation, and Unemployment)

The French national health insurance system reports the patient's place of residence with a unique geographic code. If the place of residence has more than 1000 inhabitants, the geographic code used is the zip code. For smaller communities, an aggregated zip code is used. Sociodemographic data for each patient's residential commune was collected and subsequently aggregated by geographic codes. In the rare instances where commune-level data was unavailable, the information was gathered, in order of preference, from the Public Establishment of Intermunicipal Cooperation (EPCI), arrondissement, or department.

To generate a socio-environmental score for each category, we used a Principal Component Analysis (PCA). The variables were scaled before computing the score. The first Principal Component was used for this calculation, except for familial hardship and educational barriers indices, where the Second Component was used for, as it provided a more accurate representation of the desired indices scores. To enhance the normality of the score distribution and ease interpretation, we applied a logarithmic transformation to each computed score. This transformation effectively rescaled the scores to fall within a range from 0, indicating a favourable environment, to 1, signifying an unfavourable environment.

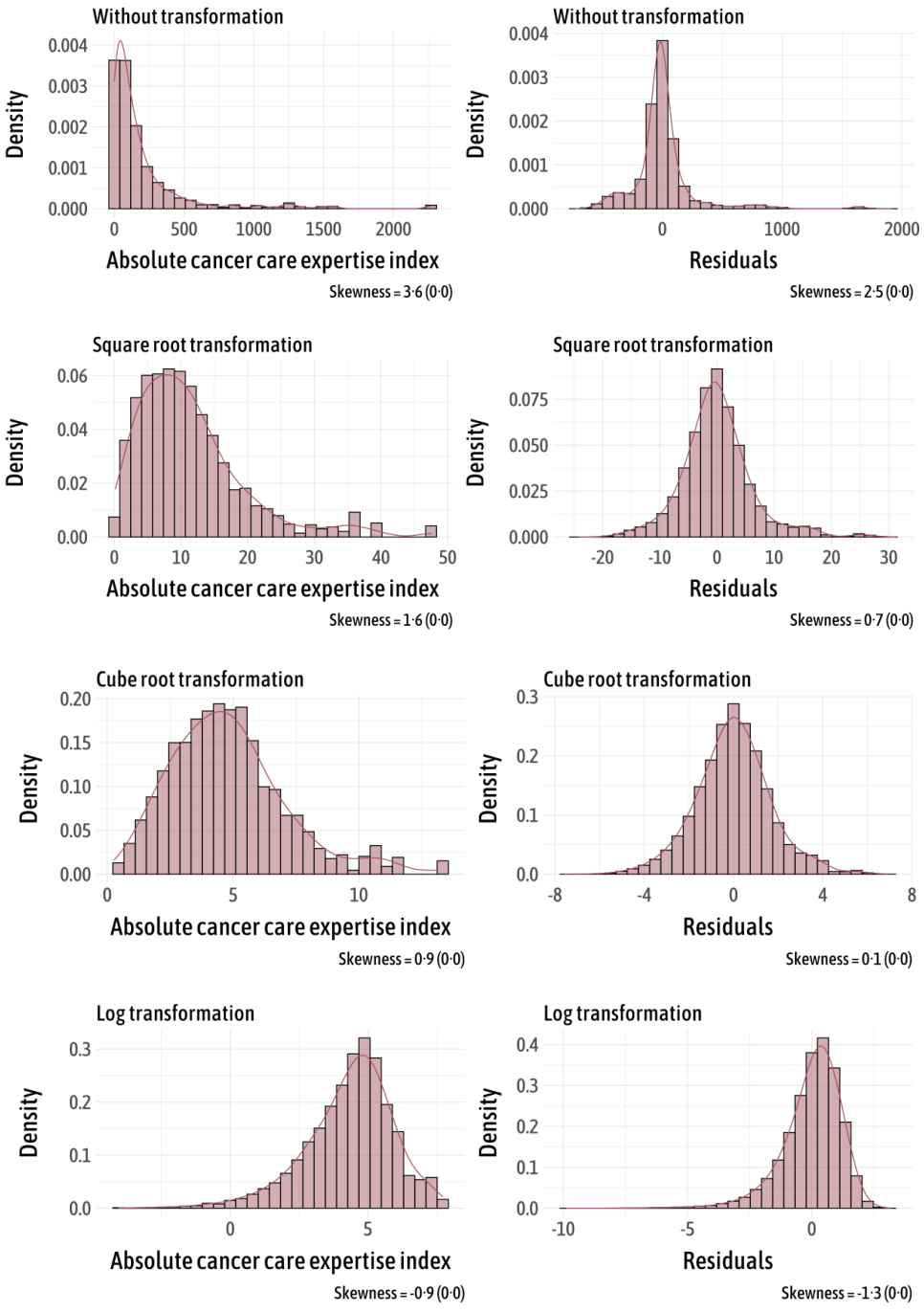



Difference in average male and female wages in total (€)
Difference in average male and female wages in intermediate occupations (€)
Average net hourly wage gap between women and men (€)
Difference in the average male and female wages of executives (€)
Difference in the average male and female wages of qualified workers (€)
Gender gap in higher education graduates
Difference in the average male and female wages of unqualified workers (€)
Difference in average male and female employee wages (€)
Difference in the number of executives between women and men
Gender gap in employment rate of 15-64 year olds
Gender segregation index of the job offer (Duncan)
Gender gap in the share of employees aged 15-64 in precarious employment
Gender gap in the share of 15-64 year olds in part-time employment
Gender
Assault and battery rate (intrafamily and non-family) per 1000 population (%)
Non-family assault and battery rate (intrafamily and non-family) per 1000 population (%)
Theft rate of vehicles per 1000 inhabitants (%)
Rate of robbery without a weapon per 1000 inhabitants (%)
Rate of thefts from vehicles per 1000 inhabitants (%)
Rate of sexual violence per 1000 inhabitants (%)
Non-violent theft rate per 1000 population (%)
Intrafamily assault and battery rate per 1000 population (%)
Rate of theft of vehicle accessories per 1000 inhabitants (%)
Rate of robbery with a weapon per 1000 inhabitants (%)
Burglary rate in dwellings per 1000 inhabitants (%)
Homicide rate per 1000 inhabitants (%)
Insecurity
Share of people aged 15 or more who are widows or widowers (%)
Share of people 80 years or older living alone (%)
Isolation
Nearest train station
Nearest subway, tramway, RER, or Transilien stop
Distance to the nearest bus stop
Transport
Retirement rate of 15 and 64 year olds (%)
Inactive rate of 15 and 64 year olds (%)
Unemployment rate of 15 and 64 year olds (%)
Employment rate of 15 and 64 year olds (%)
Average hourly net earnings of blue-collar workers (€)
Hourly wage difference between a worker and an executive (€)
Hourly wage difference between an employee and an executive (€)
Average hourly net earnings of 18 to 25 year olds (€)
Average hourly net salary of intermediate professions (€)
Average hourly net salary for employees (€)
Average hourly earnings for people over 50 (€)
Average hourly net salary for managers, senior professionals and company directors (€)
Rate of employed manual workers aged 15 to 64 (%)
Average hourly earnings for 26 to 50 year olds (€)
Unemployment

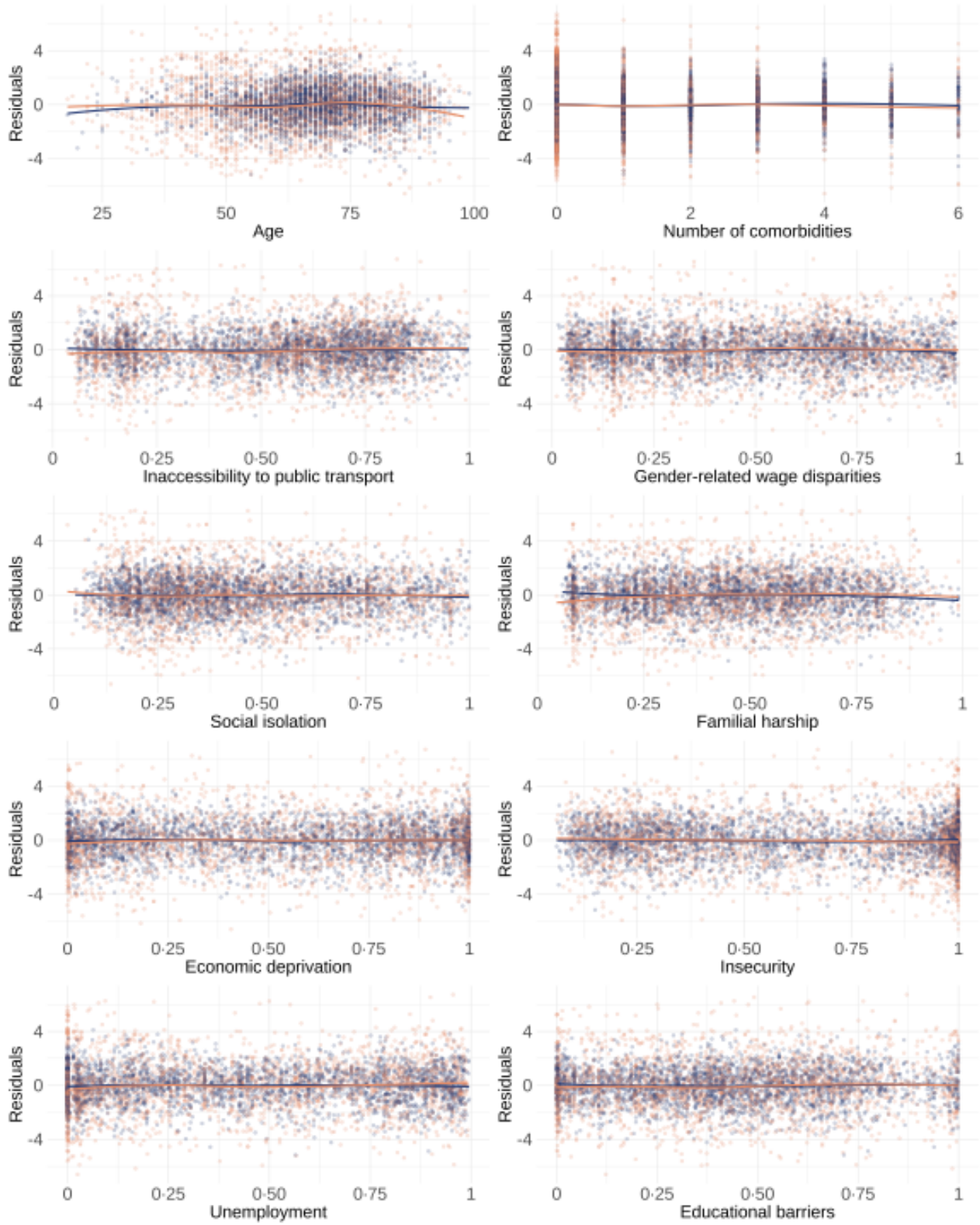
Supplementary material 6. Model assumptions. (A) Distribution of absolute cancer care expertise index: without transformation, with square root transformation, with cube root transformation, and with log transformation. (B) Residual plots for each predictor to assess linearity assumption in the absolute cancer care expertise index model with cube root transformation. (C) Distribution of relative cancer care expertise index: without transformation, with square root transformation, with cube root transformation, and with log transformation. (D) Residual plots for each predictor to assess linearity assumption in the relative cancer care expertise index model with cube root transformation.

The absolute cancer care expertise index has a right-skewed distribution with a skewness of 3.6 (0·0). The mean of absolute cancer care expertise index is 197 (316) and the median is 94 [32, 212]. Without transformation, the residuals have a right-skewed distribution of 2·5 (0·0). To better address the skewness and optimize the distribution of absolute cancer care expertise index, we considered square root transformation, cube root transformation, and log transformation. The cube root transformation proved to be more effective, bringing the skewness of residual distribution to 0·1 (0·0).

A.



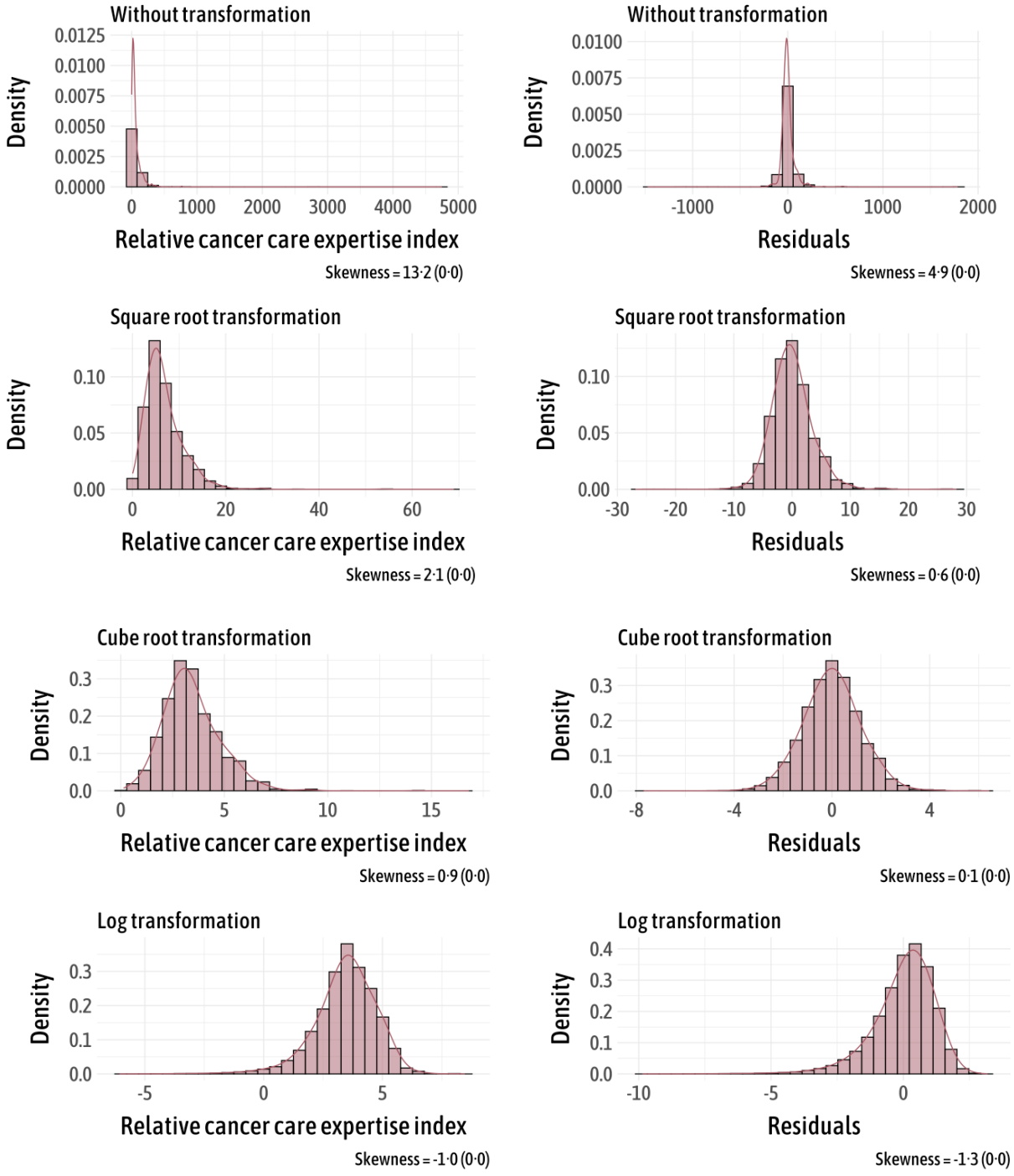
B.


Residuals
4
0
-4
25
50
75
100
Age
0
2
4
6
Number of comorbidities
0
0·25
0·50
0·75
1
Inaccessibility to public transport
Gender-related wage disparities
Social isolation
Familial harship
Economic deprivation
Insecurity
Unemployment
Educational barriers

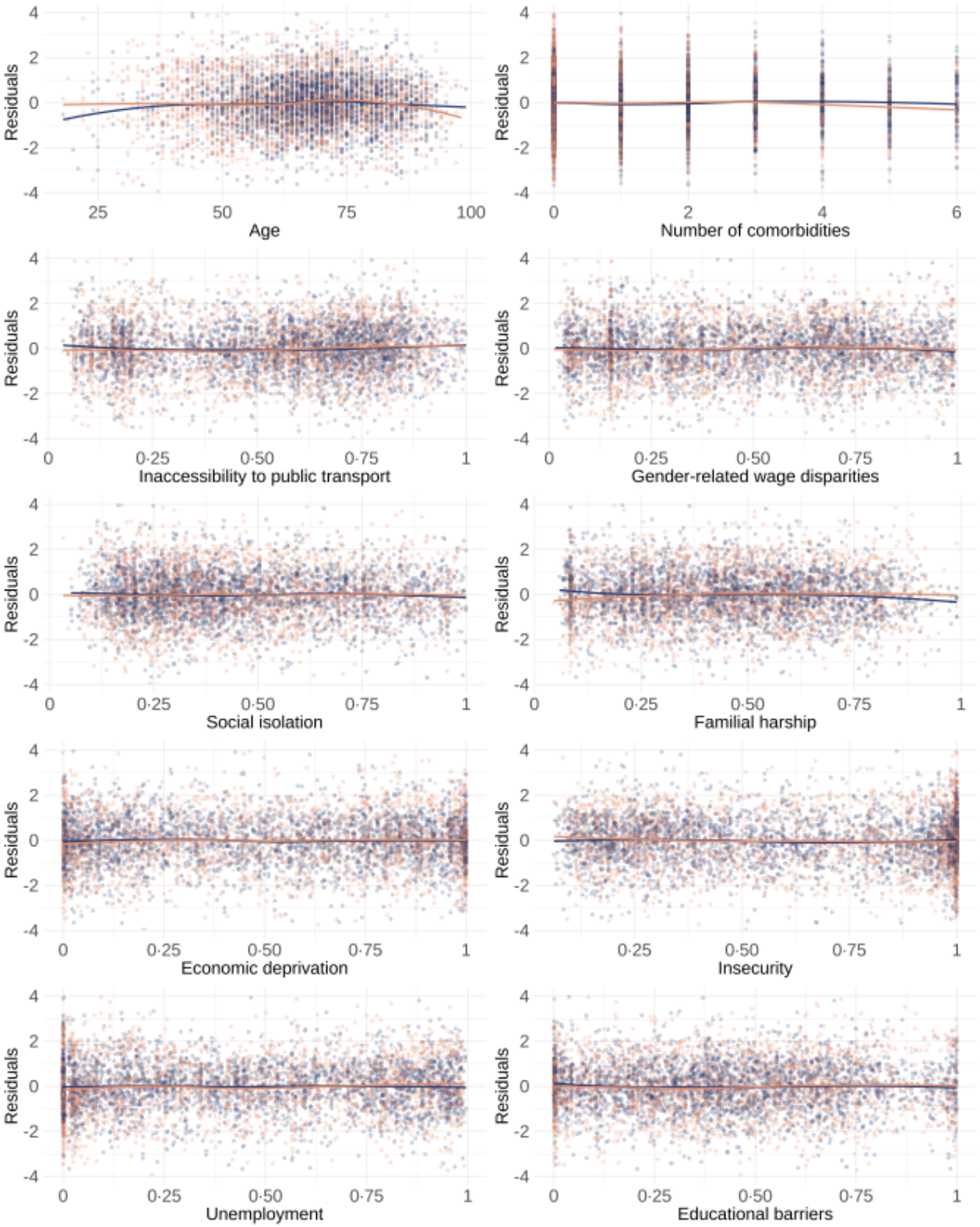
The relative cancer care expertise index is equal to the relative specific volume multiplicated by the oncological and surgical expertise score. Relative specific volume was calculated by dividing the number of patients treated for a particular cancer type in one centre by the total number of patients treated for that same cancer type across all centres, then multiplying the result by 10 000. Thus, a centre with an expertise score of 0·9 treating 20 peritoneal cancers out of 600 total peritoneal cancers this year will have a score of 20/600 * 10 000 * 0·9 = 300, and a centre with an expertise score of 0·9 treating 600 lung cancers out of 30 000 will have a score of 600/30000 * 10 000 * 0·9 = 180.

The relative cancer care expertise index has a right-skewed distribution with a skewness of 13·2 (0.0). The mean of relative cancer care expertise index is 61 (103) and the median is 34 [16, 73]. Without transformation, the residuals have a right-skewed distribution of 4·9 (0·0). To better address the skewness and optimize the distribution of relative cancer care expertise index, we considered square root transformation, cube root transformation, and log transformation. The cube root transformation proved to be again the more effective, bringing the skewness of residual distribution to 0·1 (0·0).

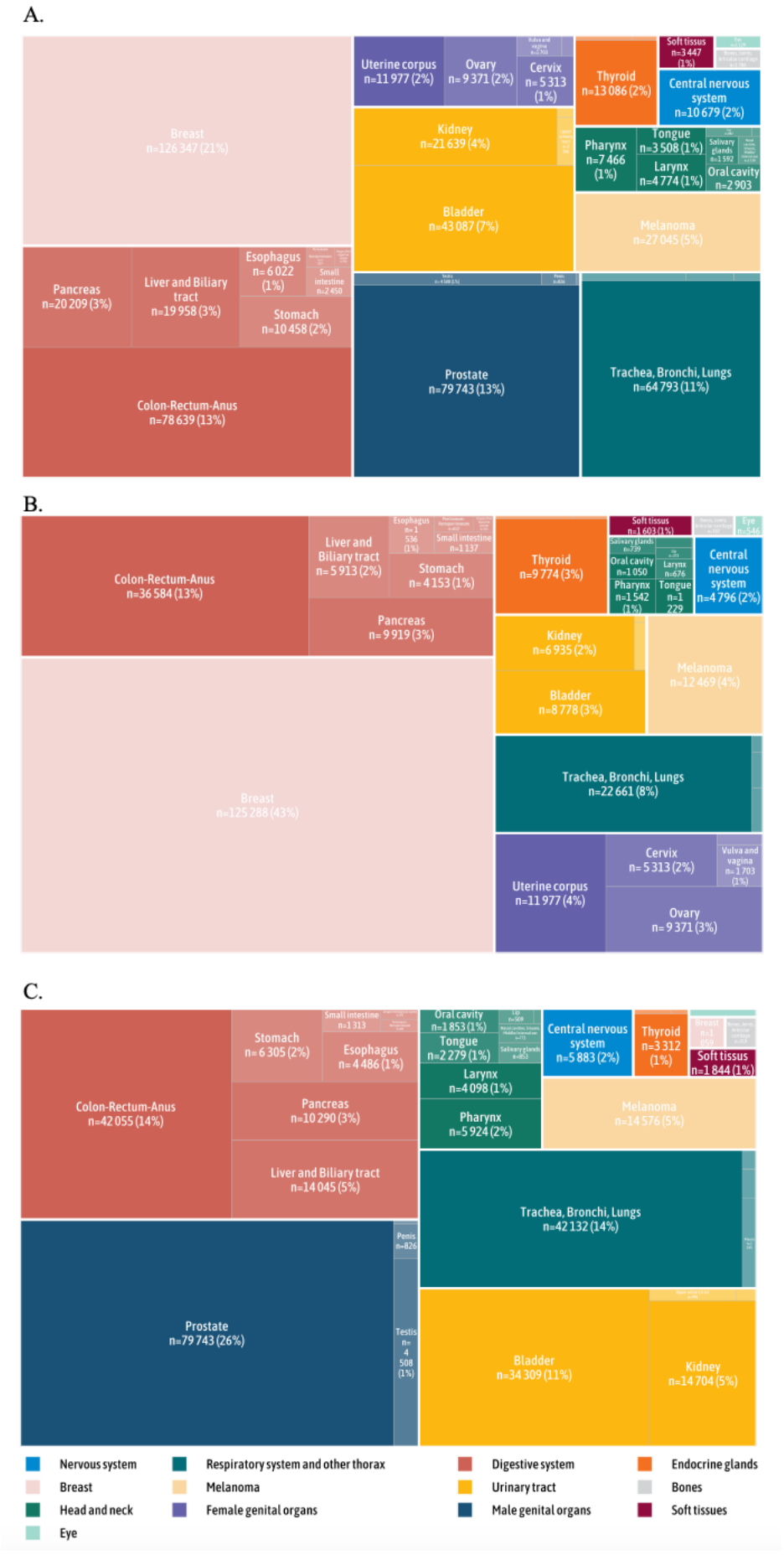
C.



D.


Residuals
Age
Number of comorbidities
Inaccessibility to public transport
Gender-related wage disparities
Social isolation
Familial harship
Economic deprivation
Insecurity
Unemployment
Educational barriers

Supplementary material 7. Distribution of cancer type. (A) Overall population. (B) Women. (C) Men.

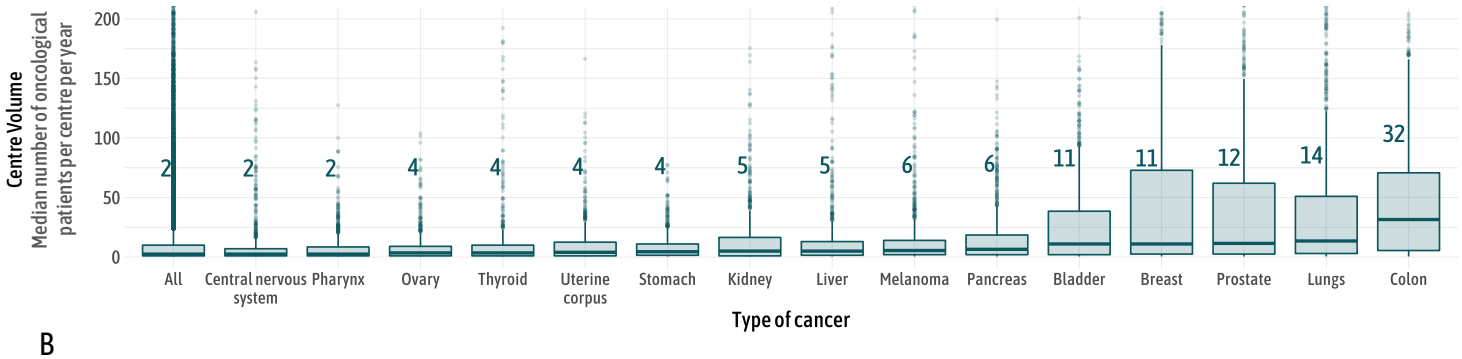


Supplementary material 8. Overview of the facilities included in the study. (A) Distribution of annual volume by cancer type, defined by the annual number of treated cancer patients per centre per cancer type. (B) Distribution of the median centre volume in which women/men are treated for their cancer. (C) Percentage of centres engaged in oncological-related activities. (D) Distribution of oncological and surgical expertise scores among centres. (E) Distribution of the distance in kilometres from patients' homes to hospitals, stratified by cancer type and gender. (F) Distribution of absolute cancer care expertise index across centres. (G) Distribution of relative cancer care expertise index across centres.

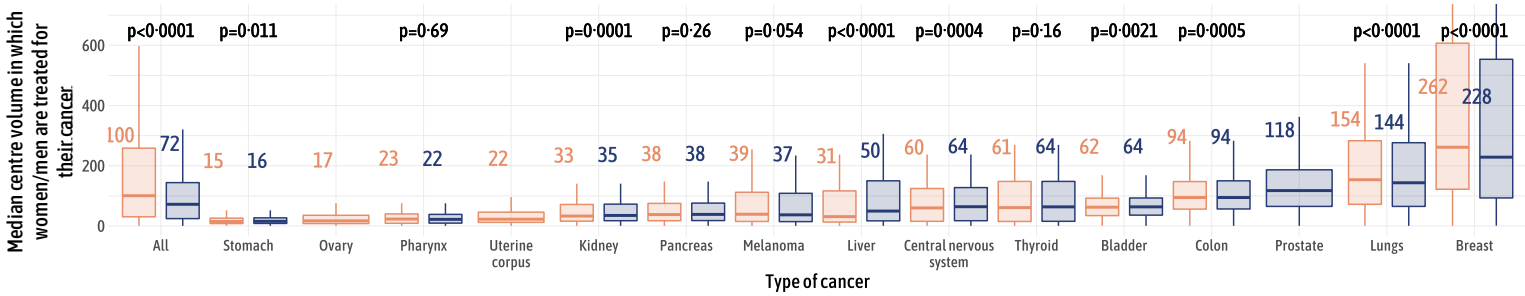
The analysis of centre volume reveals a majority of low-volume centres across all cancer types. However, for certain cancers like central nervous system cancers (with a median centre volume of 2), these low-volume centres coexist with high-volume centres that handle the majority of patients. As a result, women and men are most often treated for their central nervous system cancers at centres with median volumes of 62. In contrast, some cancer types, such as pharyngeal cancer (median centre volume at 2), exhibit a low median centre volume with only a few high-volume centres. Consequently, women and men with pharyngeal cancer typically received treatment at centres with a median annual volume of 23.

In Figures A, B, and E, boxplots are used to depict the data distribution. They display the median and interquartile range (IQR). Whiskers on these plots extend up to 1.5 times the IQR from the box. Figure A also shows individual data points beyond the whiskers, indicating potential outliers. The p-value for Figure B signifies the statistical significance of differences between the two gender groups concerning centre volume, whereas for Figure E, it indicates the significance of differences related to distance.

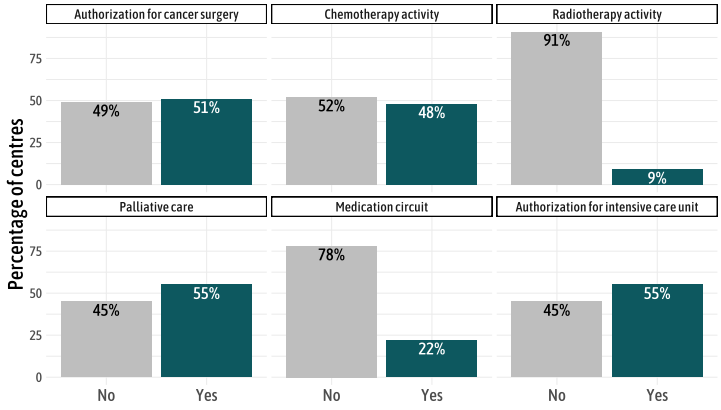
A



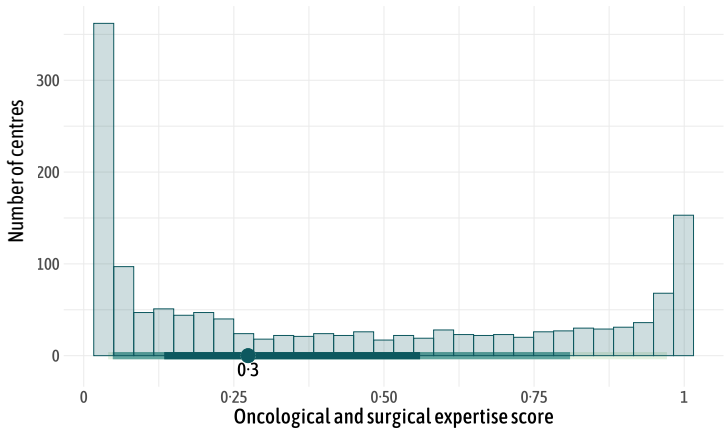
B



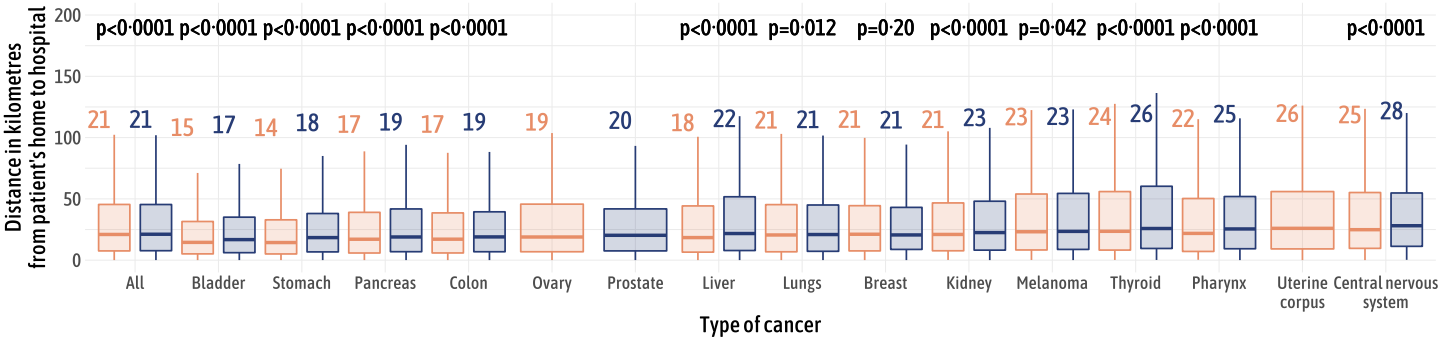
C



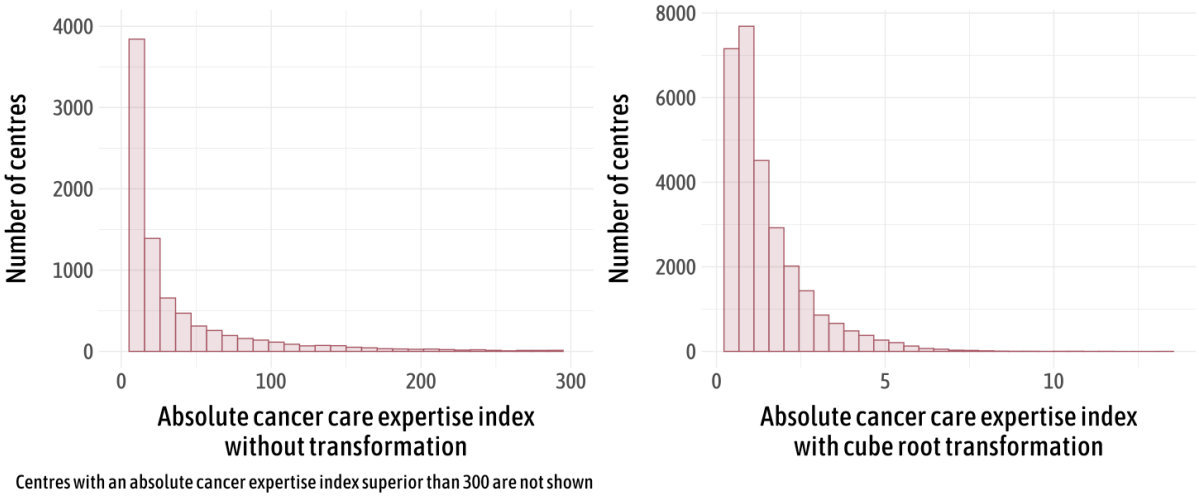
D



E



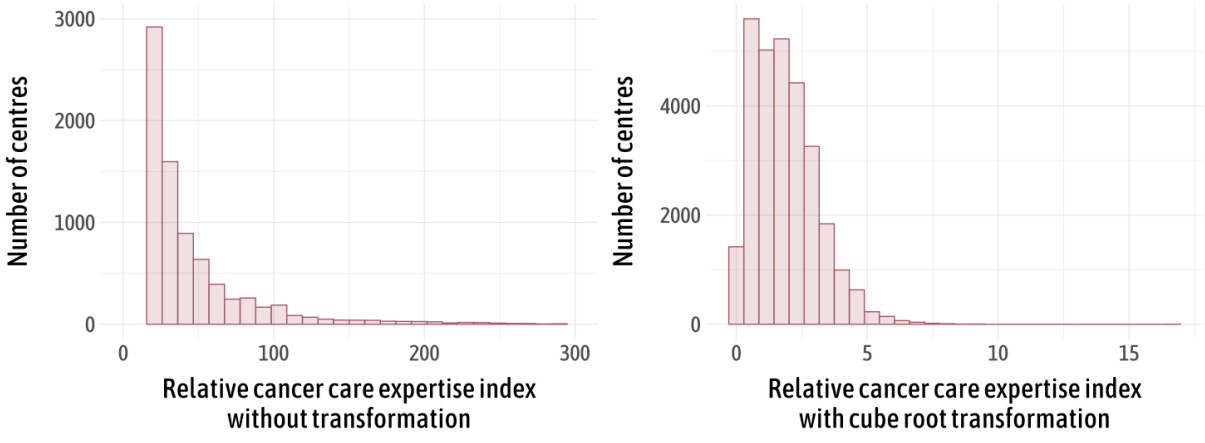
F



Centres with an absolute cancer expertise index superior than 300 are not shown

A centre might appear multiple times in this plot due to variations in its absolute cancer care expertise index across different cancer types.

G

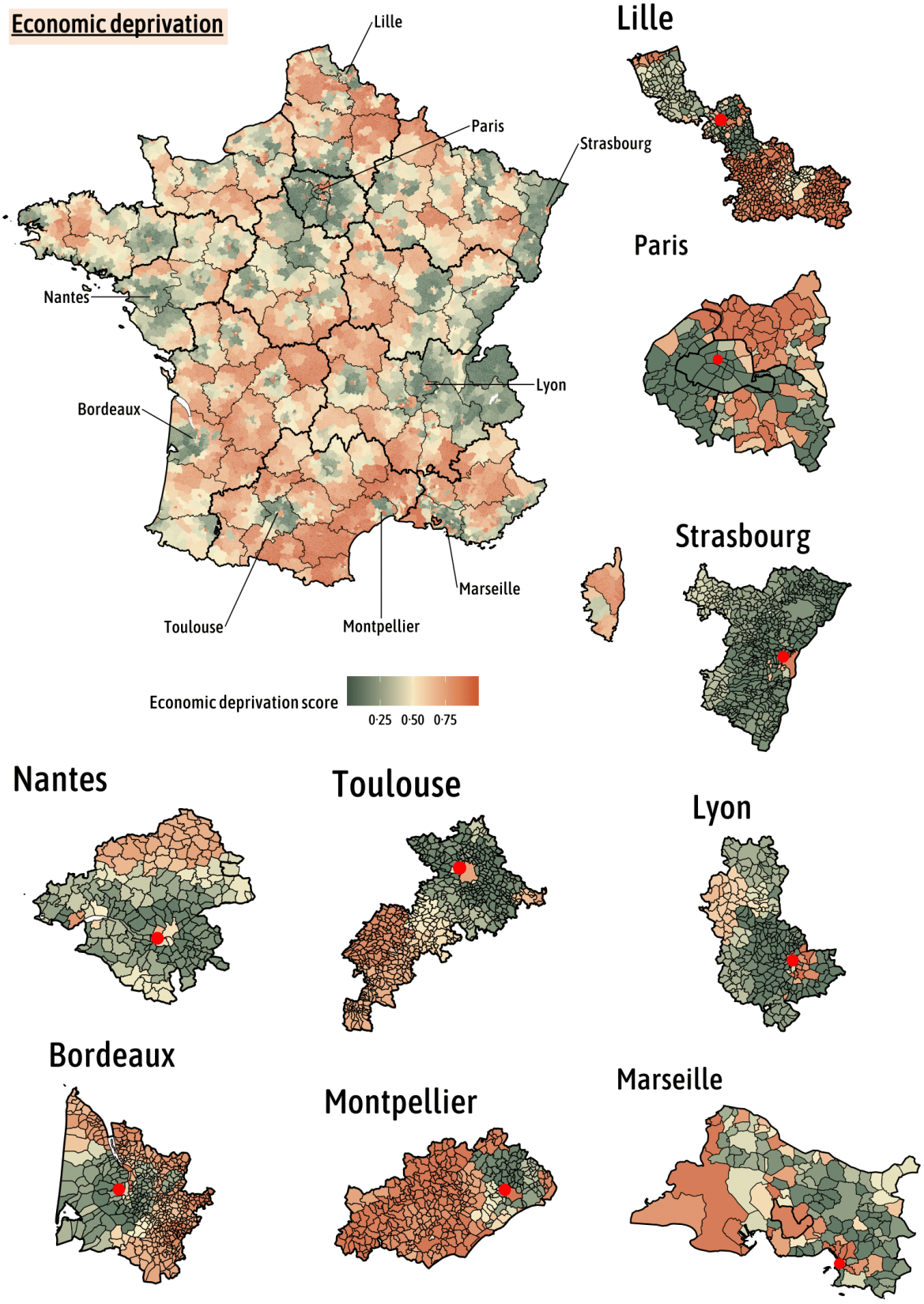


Centres with a relative cancer expertise index superior than 300 are not shown

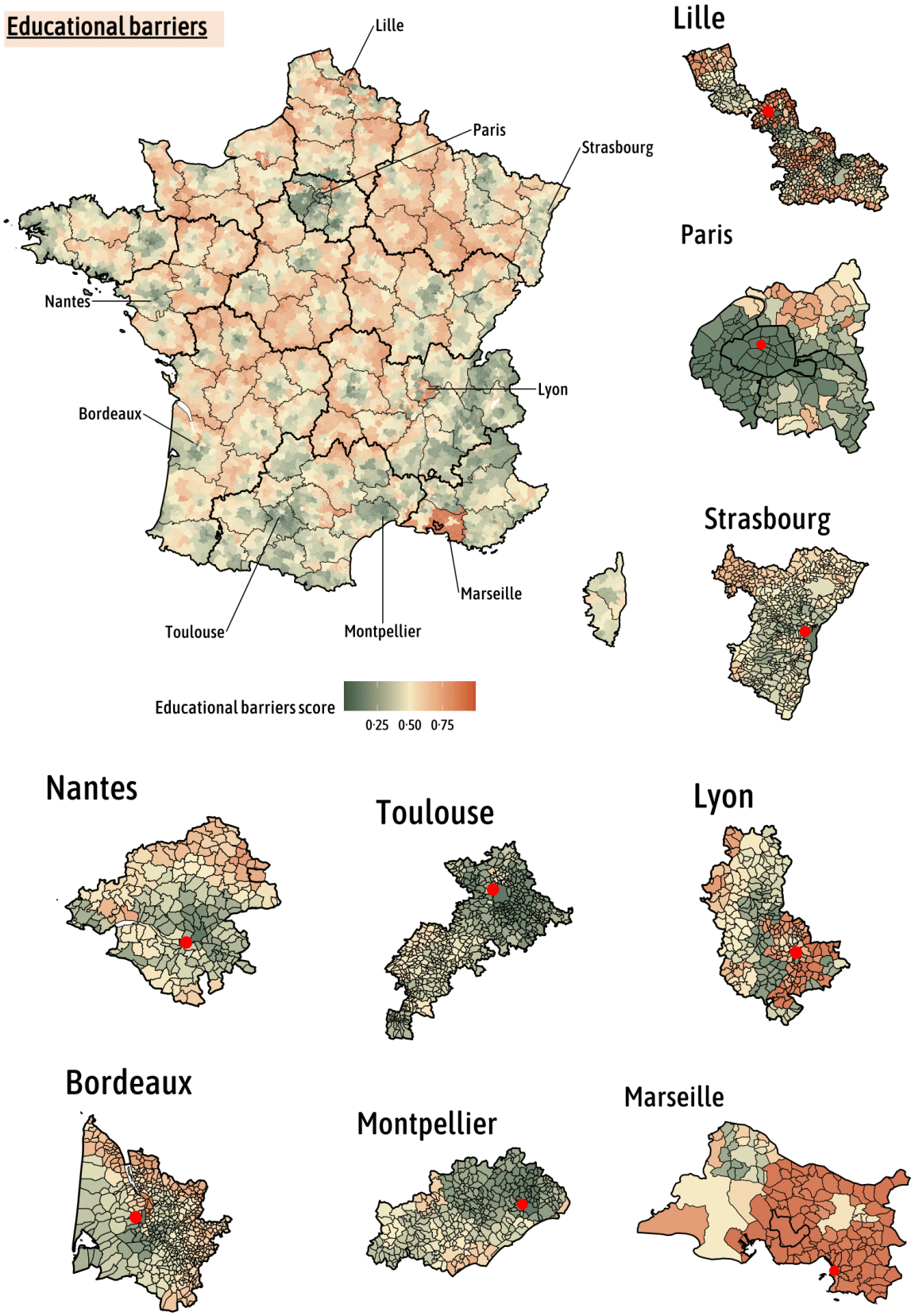
A centre might appear multiple times in this plot due to variations in its relative cancer care expertise index across different cancer types.

Supplementary material 9. Mapping of socio-environmental indices. (A) Economic deprivation. (B) Educational barriers. (C) Gender-related wage disparities. (D) Unemployment. (E) Social isolation. (F) Inaccessibility to public transport. (G) Insecurity. (F) Familial hardship.

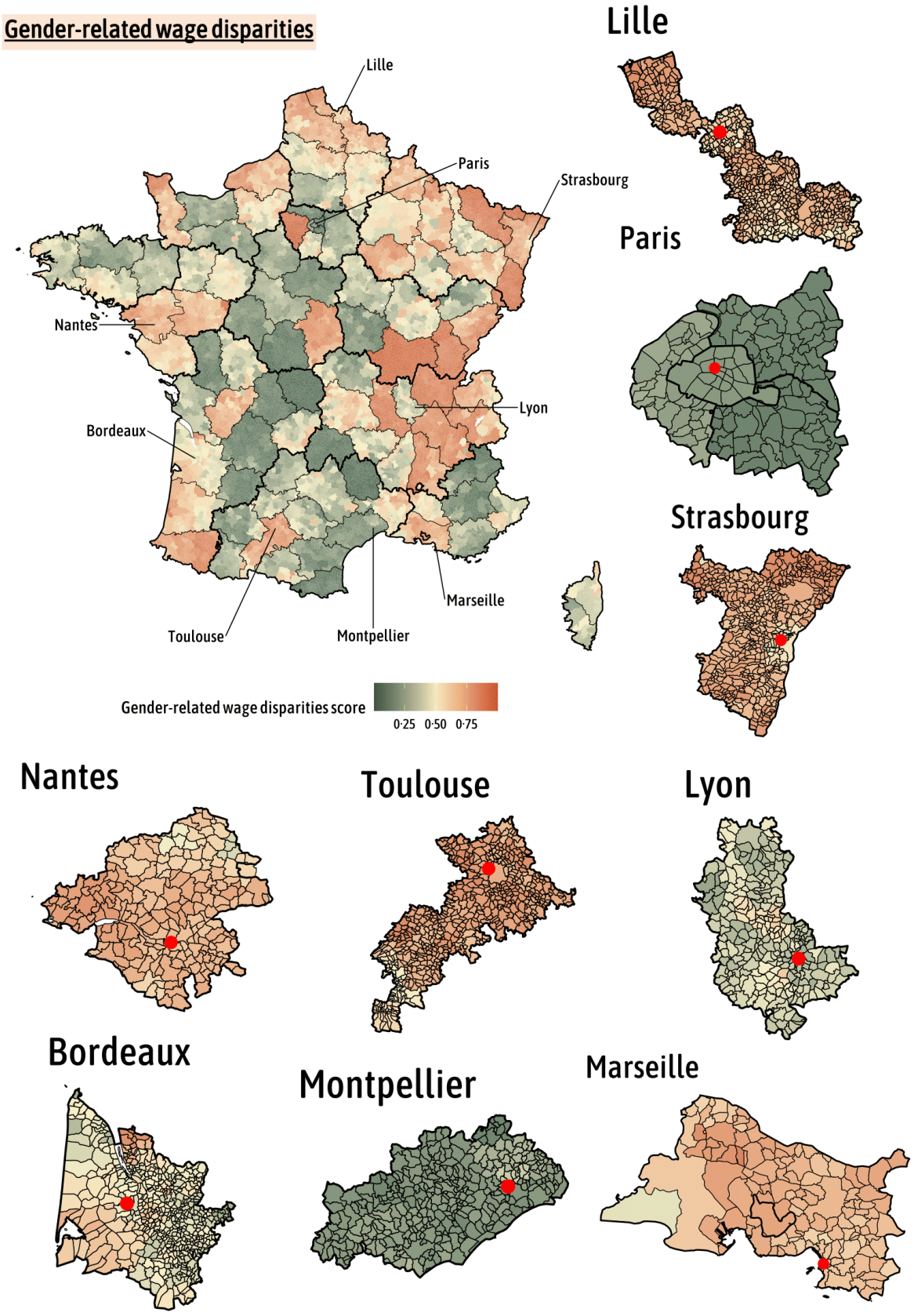
A.



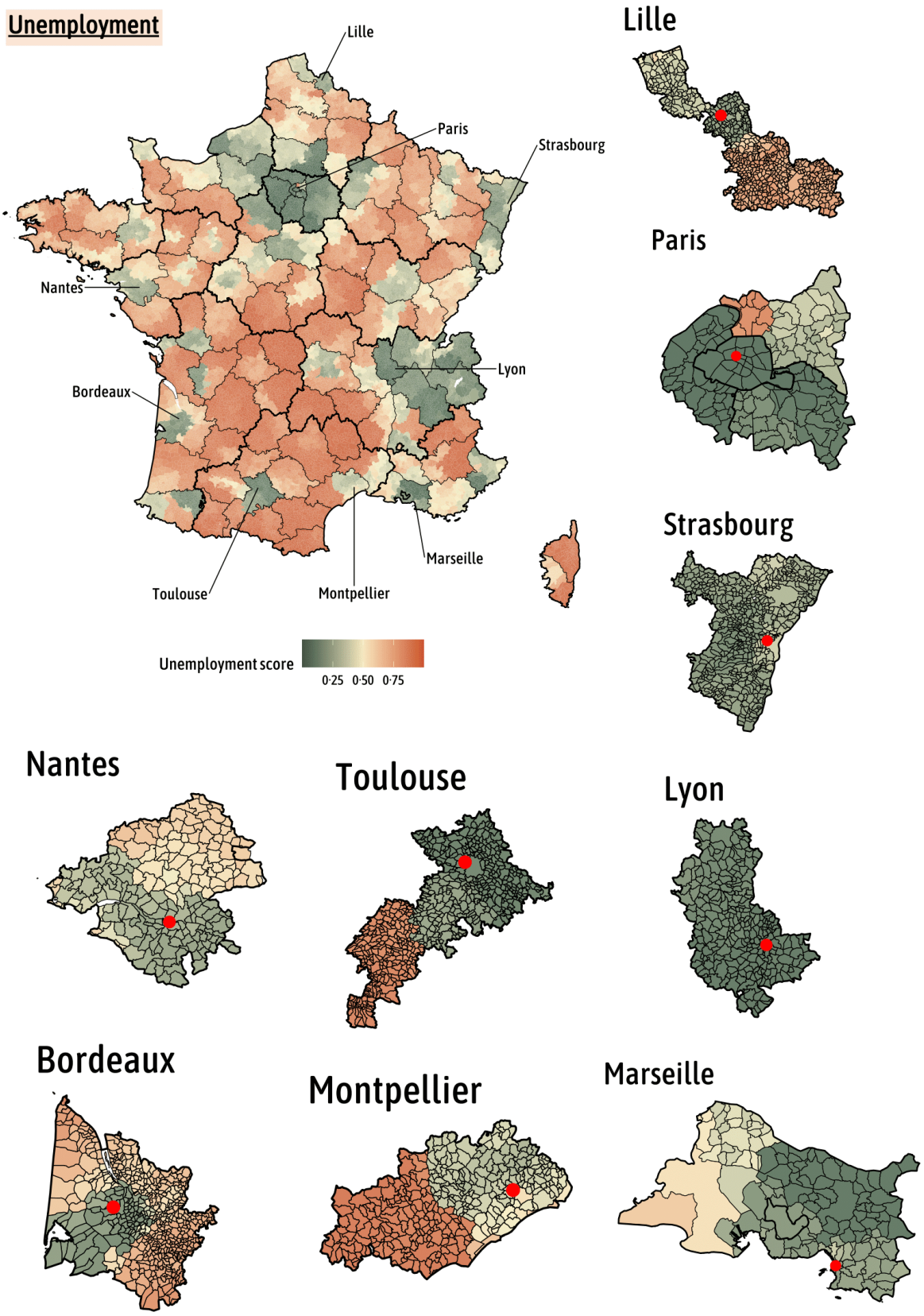
B.


Educational barriers
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Toulouse
Montpellier
Marseille
Educational barriers score
0·25 0·50 0·75
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

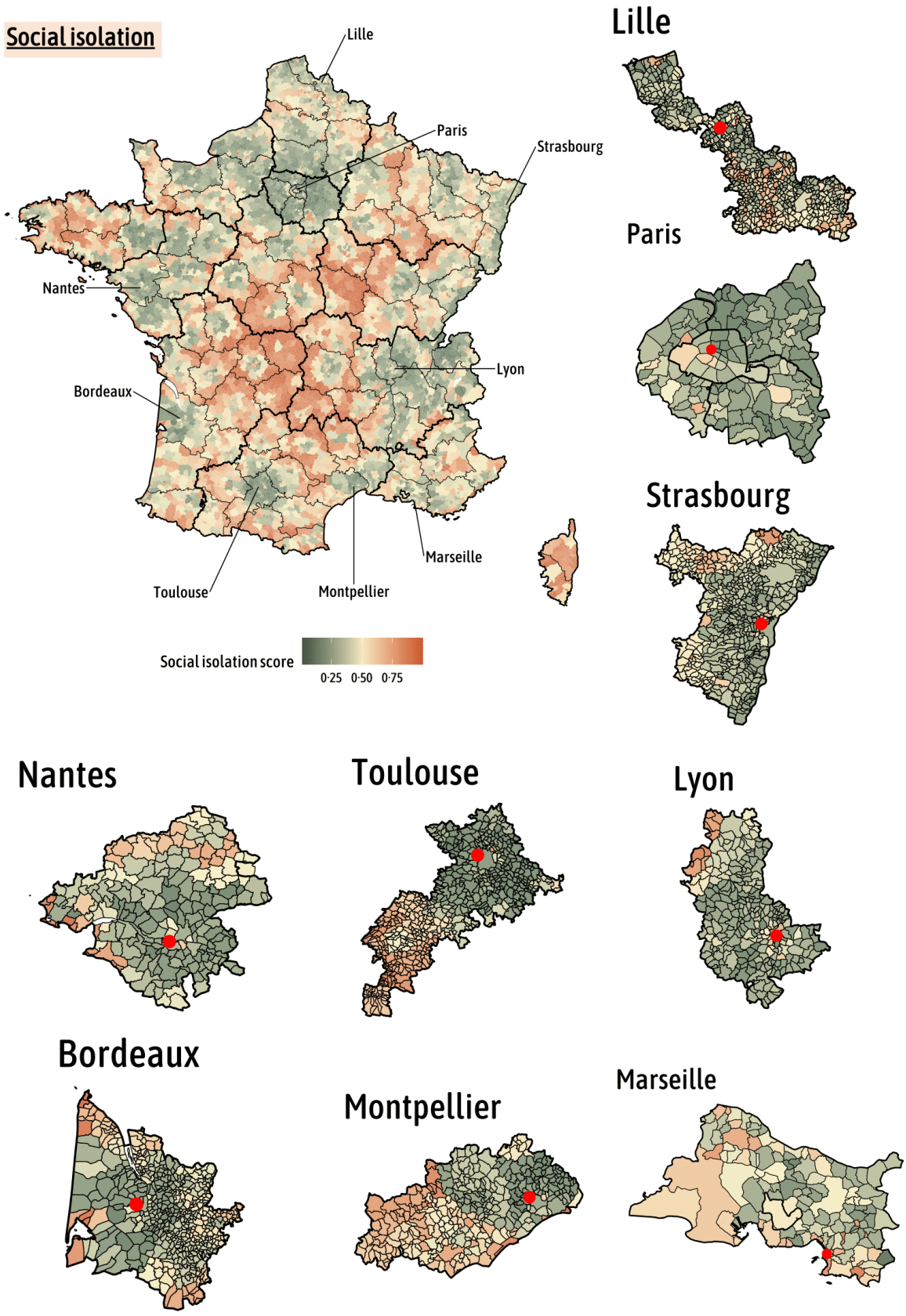
C.


Gender-related wage disparities
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Marseille
Toulouse
Montpellier
Gender-related wage disparities score
0·25 0·50 0·75
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

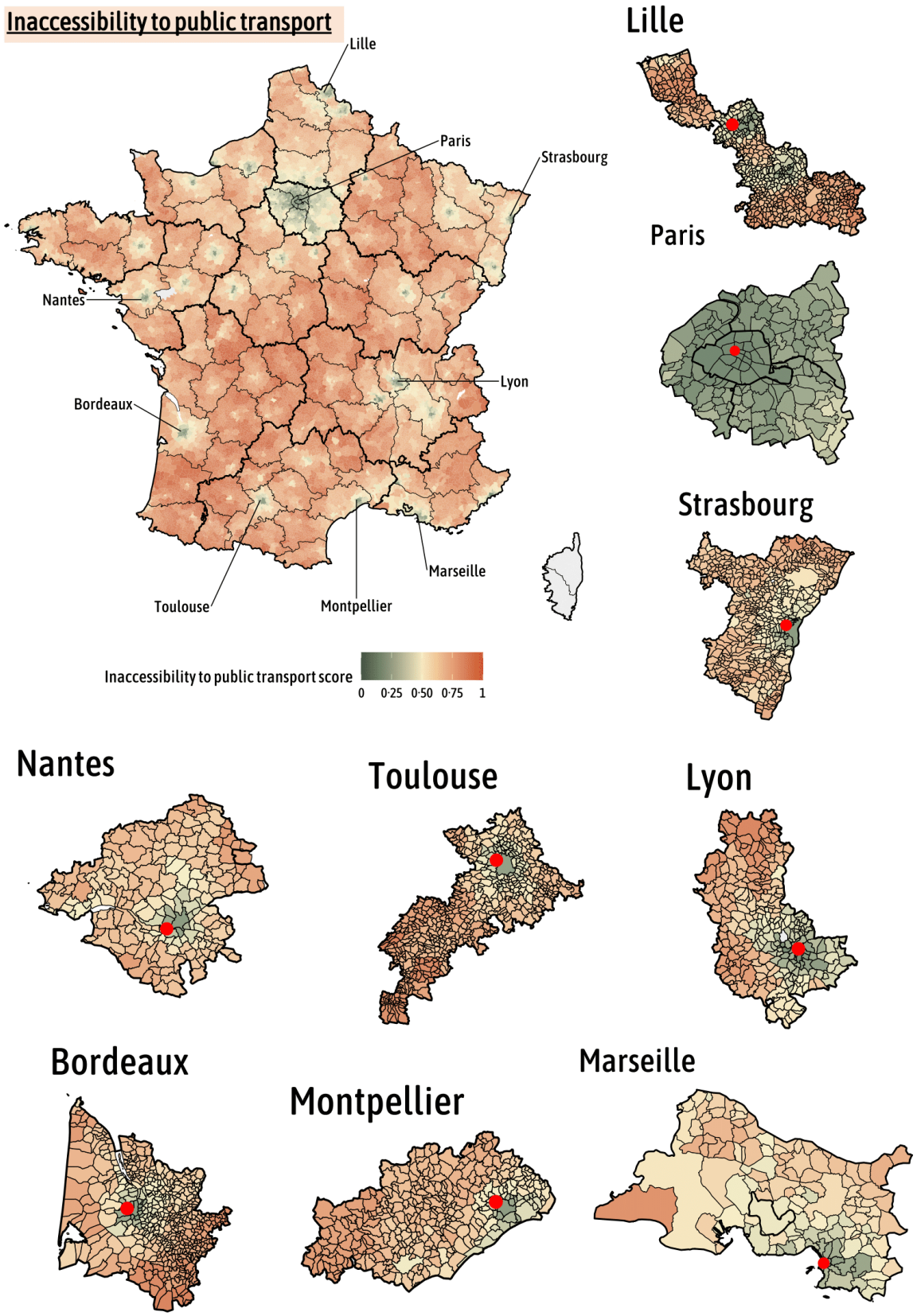
D.


Unemployment
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Marseille
Toulouse
Montpellier
Unemployment score
0·25 0·50 0·75
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

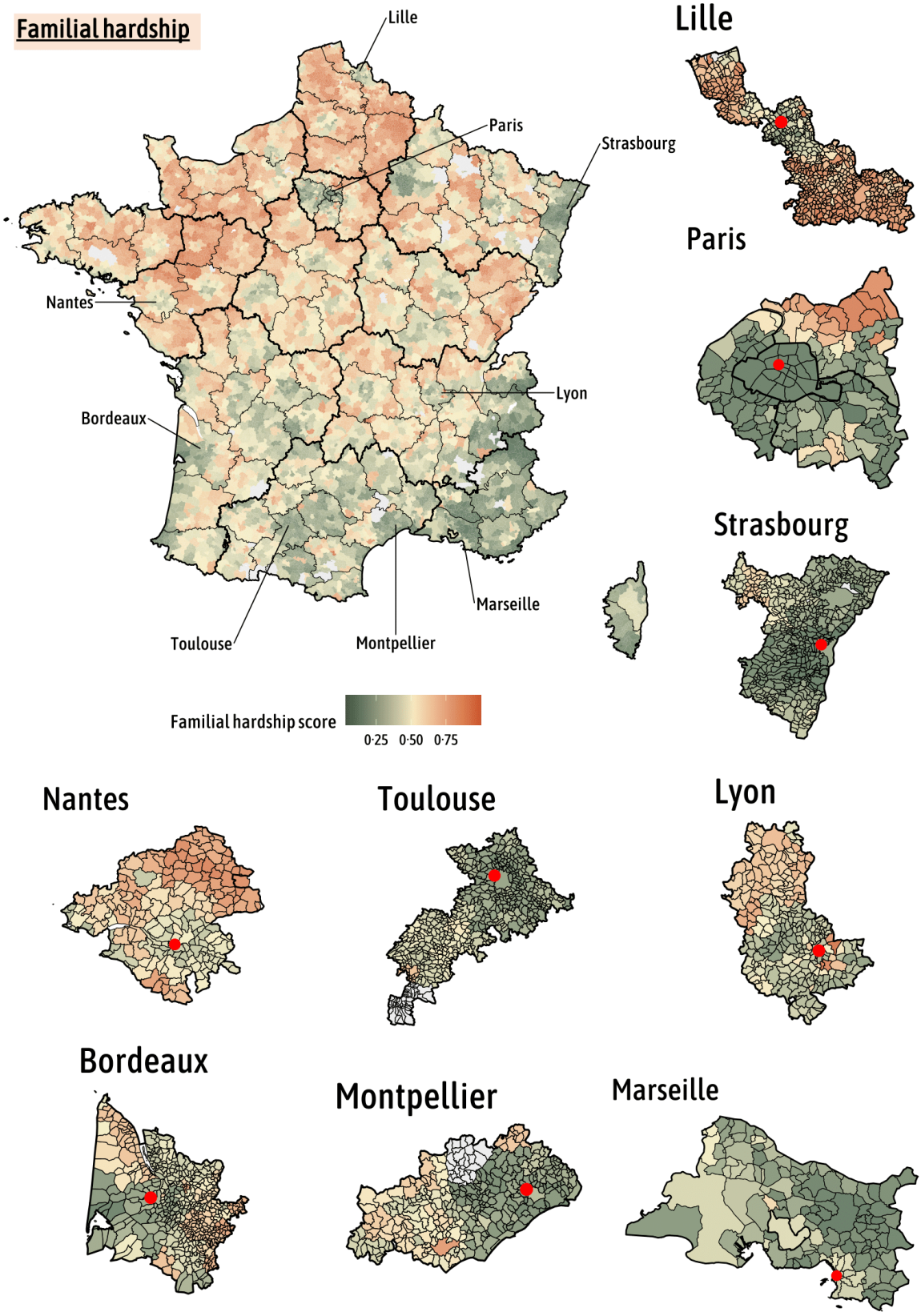
E.


Social isolation
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Marseille
Toulouse
Montpellier
Social isolation score
0·25 0·50 0·75
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

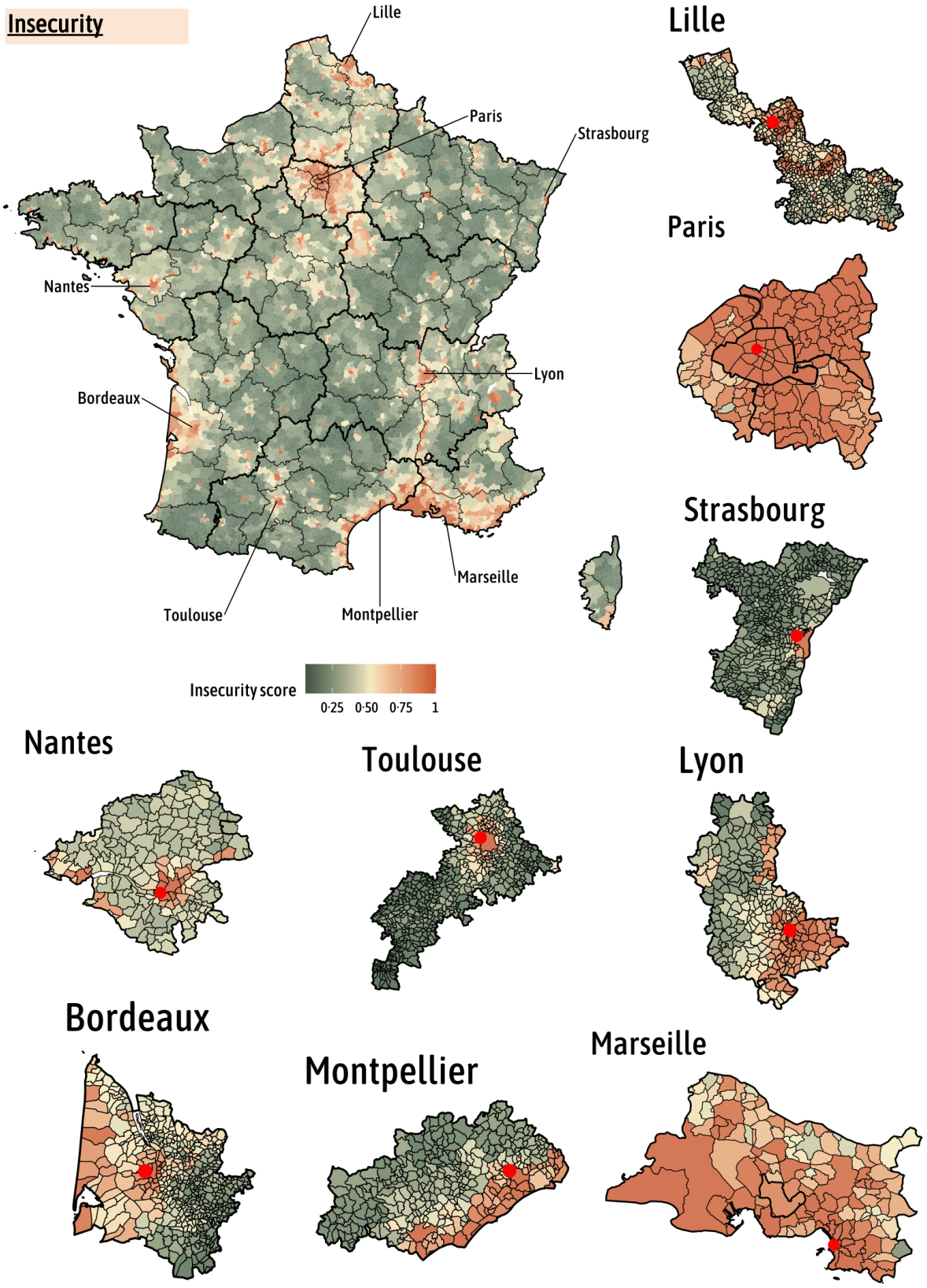
F.


Inaccessibility to public transport
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Marseille
Toulouse
Montpellier
Inaccessibility to public transport score
0 0·25 0·50 0·75 1
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

G.


Familial hardship
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Marseille
Toulouse
Montpellier
Familial hardship score
0·25 0·50 0·75
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

H.


Insecurity
Lille
Paris
Strasbourg
Nantes
Lyon
Bordeaux
Toulouse
Montpellier
Marseille
Insecurity score
0·25 0·50 0·75 1
Lille
Paris
Strasbourg
Nantes
Toulouse
Lyon
Bordeaux
Montpellier
Marseille

Supplementary material 10. Sensitivity analysis: Multivariable mixed linear model with only non-sex-specific cancers

We used for this sensitivity analysis the absolute cancer care expertise index and relative cancer care expertise index after cube root transformation.

	Multivariable mixed linear regression model with only non-sex specific cancers for absolute cancer care expertise index				Multivariable mixed linear regression model with only non-sex specific cancers for relative cancer care expertise index			
	Coefficient	95% CI	p-value*	Interaction Men-Women**	Coefficient	95% CI	p-value	Interaction Men-Women**
Gender				p<0·0001				p<0·0001
Men	4·20	[3·70, 4·71]	p<0·0001		5·74	[5·16, 6·32]	p<0·0001	
Women	5·36	[4·77, 5·95]	p<0·0001		6·07	[5·43, 6·71]	p<0·0001	
Age, years				p<0·0001				0·29
Men	-0·01	[-0·01, -0·01]	p<0·0001		-0·01	[-0·01, -0·01]	p<0·0001	
Women	-0·02	[-0·02, -0·02]	p<0·0001		-0·01	[-0·01, -0·01]	p<0·0001	
Number of comorbidities				p<0·0001				p<0·0001
Men	0·00	[0·00, 0·01]	0·013		0·01	[0·00, 0·01]	p<0·0001	
Women	-0·05	[-0·06, -0·04]	p<0·0001		-0·03	[-0·03, -0·02]	p<0·0001	
Inaccessibility to public transport				p<0·0001				p<0·0001
Men	-0·53	[-0·59, -0·48]	p<0·0001		-0·41	[-0·46, -0·37]	p<0·0001	
Women	-1·20	[-1·33, -1·06]	p<0·0001		-0·69	[-0·79, -0·6]	p<0·0001	
Familial hardship				p<0·0001				0·055
Men	-0·53	[-0·57, -0·49]	p<0·0001		-0·39	[-0·42, -0·36]	p<0·0001	
Women	-0·69	[-0·80, -0·59]	p<0·0001		-0·43	[-0·50, -0·36]	p<0·0001	
Social isolation				p<0·0001				p<0·0001
Men	-0·13	[-0·17, -0·09]	p<0·0001		-0·11	[-0·14, -0·08]	p<0·0001	
Women	-0·40	[-0·49, -0·31]	p<0·0001		-0·24	[-0·31, -0·18]	p<0·0001	
Insecurity				p<0·0001				p<0·0001
Men	-0·06	[-0·10, -0·01]	0·0075		-0·03	[-0·06, 0·00]	0·041	
Women	-0·32	[-0·41, -0·22]	p<0·0001		-0·14	[-0·21, -0·08]	p<0·0001	
Economic deprivation				p<0·0001				0·0084
Men	-0·05	[-0·08, -0·01]	0·0044		-0·04	[-0·06, -0·02]	0·0003	
Women	-0·14	[-0·21, -0·06]	p<0·0001		-0·08	[-0·13, -0·03]	p<0·0001	
Educational barriers				0·0001				0·039
Men	-0·03	[-0·06, 0·01]	0·15		-0·01	[-0·03, 0·02]	0·50	
Women	-0·12	[-0·20, -0·04]	p<0·0001		-0·04	[-0·10, 0·01]	0·039	
Gender-related wage disparities				p<0·0001				p<0·0001
Men	-0·01	[-0·04, 0·02]	0·47		-0·01	[-0·03, 0·01]	0·16	
Women	-0·11	[-0·18, -0·04]	p<0·0001		-0·07	[-0·12, -0·02]	p<0·0001	
Unemployment				0·0002				0·0006
Men	-0·02	[-0·05, 0·02]	0·44		-0·04	[-0·07, -0·01]	0·0048	
Women	0·09	[0·00, 0·18]	0·0091		0·03	[-0·04, 0·09]	0·27	

CI: Confidence interval

* The p value demonstrates the significance of the relationships between the given social determinant and the cancer care expert index for each gender separately.

** The p value indicates the significance of the interaction between men and women for each factor on the cancer care expertise indices

REFERENCES

1 Bannay A, Chaignot C, Blotière P-O, *et al.* The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. *Med Care* 2016; **54**: 188–94.

2 Billioti de Gage S, Collin C, Le-Tri T, *et al.* Antidepressants and Hepatotoxicity: A Cohort Study among 5 Million Individuals Registered in the French National Health Insurance Database. *CNS Drugs* 2018; **32**: 673–84.

3 Bouyer B, Rudnichi A, Dray-Spira R, Zureik M, Coste J. Thromboembolic risk after lumbar spine surgery: a cohort study on 325 000 French patients. *J Thromb Haemost* 2018; published online June 12. DOI:10.1111/jth.14205.

4 Coste J, Blotiere P-O, Miranda S, *et al.* Risk of early neurodevelopmental disorders associated with in utero exposure to valproate and other antiepileptic drugs: a nationwide cohort study in France. *Sci Rep* 2020; **10**: 17362.

5 Giral P, Neumann A, Weill A, Coste J. Cardiovascular effect of discontinuing statins for primary prevention at the age of 75 years: a nationwide population-based cohort study in France. *Eur Heart J* 2019; **40**: 3516–25.

6 Jabagi MJ, Vey N, Goncalves A, Tri TL, Zureik M, Dray-spira R. Risk of secondary hematologic malignancies associated with breast cancer chemotherapy and G-CSF support: A nationwide population-based cohort. *International Journal of Cancer* 2020; **n/a**. DOI:https://doi.org/10.1002/ijc.33216.

7 Lemaitre M, Kirchgesner J, Rudnichi A, *et al.* Association Between Use of Thiopurines or Tumor Necrosis Factor Antagonists Alone or in Combination and Risk of Lymphoma in Patients With Inflammatory Bowel Disease. *JAMA* 2017; **318**: 1679–86.

8 Maura G, Bardou M, Billionnet C, Weill A, Drouin J, Neumann A. Oral anticoagulants and risk of acute liver injury in patients with nonvalvular atrial fibrillation: a propensity-weighted nationwide cohort study. *Scientific Reports* 2020; **10**: 11624.

9 Sbidian E, Mezzarobba M, Weill A, Coste J, Rudant J. Persistence of treatment with biologics for patients with psoriasis: a real-world analysis of 16 545 biologic-naïve patients from the French National Health Insurance database (SNIIRAM). *Br J Dermatol* 2019; **180**: 86–93.

10 Semenzato L, Botton J, Drouin J, Cuenot F, Weill A, Zureik M. Maladies chroniques, états de santé et risque d'hospitalisation et de décès hospitalier pour COVID-19 lors de la première vague de l'épidémie en France: Étude de cohorte de 66 millions de personnes. 2021; : 41.

11 World Health Organization. International Classification of Diseases (ICD). https://www.who.int/standards/classifications/classification-of-diseases.