WASABY

WAter and Soil contamination and Awareness on Breast cancer risk in Young women

MD.1

SCIENTIFIC INTERIM REPORT

From 1.1.2018 to 30.6.2019

Project Leader: Paolo Baili
Working Team: Roberto Lillini, Camilla Amati, Alessandro Borgini, Martina Bertoldi, Andrea Tittarelli
Representatives of the Associated Partners: David Ritchie, Alexander Katalinic, Ron Pritzkuleit, Guy Launoy, Ludivine Launay, Elodie Guillaume, Tina Žagar, Giovanna Tagliabue, Paolo Contiero
Website: http://www.wasabysite.it
V2 – 10th September 2019

WASABY

INDEX

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

1. THE WASABY PROJECT

1.1. INTRODUCTION

The dissemination of knowledge on breast cancer risk factors among young women in areas with higher-than-average breast cancer incidence in EU due to water and soil environmental contamination is at the centre of call PP-2-5-2016[1], under which the WASABY project was subsidized. Under the rationale that knowing the existence and the causes of cancer incidence inequalities is a key requisite to identify measures for reducing them, WASABY proposed to address the call as follows:

1) by mapping breast cancer risk and identifying areas at higher risk via geographic information systems
2) by reviewing scientific literature on the relationship between water/soil pollutants and breast cancer risk, so as to organize a pilot environmental study;
3) by promoting primary prevention actions potentially reducing cancer risk among young women in areas at higher risk.

We proposed to design a model able to identify areas with higher cancer rates to study whether pollutant contamination may be a cause for increased cancer risk. Hence the project aim is to improve the health of EU citizens and reduce health inequalities. WASABY acts in three public health domains:

- ✓ Cancer information system: one of the aims is to improve the European population-based cancer information system by promoting models for the use of Cancer Registry (CR) data. The study of cancer risk using geographical spatial analysis is one of the major public health topics for CRs: public health managers are enabled to prioritize actions in the areas of their territories most in need
- ✓ Research for public health: performing spatial analysis on cancer risk, estimating deprivation indexes, literature review, correlation analysis between environmental pollutant and cancer incidence data are research activities performed to improve public health knowledge
- ✓ Cancer Prevention: online courses are promoted and organized to increase awareness of breast cancer risk among young girls and adolescents

This interim report is organized this way:

- ✓ Chapter 1 summarizes information included in the project application
- ✓ Chapter 2 summarizes the main activities performed in the first 18 months of the project
- ✓ Chapter 3 summarizes the main future activities
- ✓ List of annexes includes milestones and deliverables already produced and published online

[1] http://ec.europa.eu/research/participants/portal4/desktop/en/opportunities/3hp/topics/pp-2-5-2016.html

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

1.2. SPECIFIC OBJECTIVES

To accomplish the points above, the following tasks and specific objectives (SO) were foreseen:

Identification of areas with an increased breast cancer risk (SO-1). CRs are the main stakeholders in the first part of the action: as the action focuses on exposures with long latency, the use of population-based data is cost effective compared to experimental or other prospective studies. We are pursuing the following steps:

a) Identification of participating CRs
b) Collection of all breast cancer cases of participating CRs
c) Utilization of Socio Economic Status (SES) data (SO-2)
d) Identification of areas with higher-than-CR average rates

Analysing relationship between environmental data and cancer incidence following these steps:

a) Literature review on water/soil environmental risk factors for breast cancer (SO-3)
b) Literature review on methodologies to be used in ecological environmental studies (SO-4)
c) Collection of available data on pollutants across Europe contacting European and/or National Environmental Protection Agencies (SO-4)
d) A pilot environmental study on breast cancer and water/soil contaminants will be be carried out on one selected area characterised by higher than average breast cancer incidence (SO-4). The relation between pollutants and breast cancer risk in the pilot area is investigated through spatial analysis evaluated in literature review indicated in point b).

Online primary prevention courses (SO-5) as risk reducing behaviours can lead to a reduction of 50% of breast cancer incidence as indicated by Colditz [Colditz G, CA Cancer J Clin, 2014]:

a) courses able to increase awareness on breast cancer risk factors will be performed in a selected number of countries participating in the project
b) The courses will be directed to reach young girls and adolescents (age 12 to 19) as these ages fall in the critical windows of susceptibility to known effect of epigenetic marks, endocrine disruption and carcinogens
c) Courses will use an interactive, anonymous, online and-user friendly platform which facilitates the easy adoption of correct habits in an attractive way.

This interim report will describe the activities performed in the first 18 Months to reach the various SOs:

- ✓ List of CRs involved and list of CRs for which breast cancer risk maps are already performed
- ✓ List of countries for which European Deprivation Indexes are going to be estimated
- ✓ Scientific Literature Reviews performed on water/soil environmental risk factors and breast cancer
- ✓ Activities to be performed to get ready for the online courses

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

1.3. WORK PACKAGES

WP	Title	Description	Leader
1	Coordination of the project	Actions undertaken to manage the project and to make sure that it is implemented as planned	Fondazione IRCCS Istituto Nazionale dei Tumori (INT)
2	Dissemination of the project	Actions undertaken to ensure that the results and deliverables will be made available to the target groups	Association Europeenne Des Ligues Contre Le Cancer (ECL)
3	Evaluation of the project	Actions undertaken to verify if project is being implemented as planned and reaches the objectives	Universität Zu Lübeck (UZL)
4	Data management	Actions undertaken to manage and coordinate data for and from WP-5, WP-6 and WP-7	INT
5	Deprivation indexes	Construction of indicators of social and economic deprivation by the smallest geographical areas of countries	Université De Caen (UNICEN)
6	Methods and analysis	Actions undertaken to design and validate a method for studying and mapping the variation of cancer risk within CRs areas	Onkoloski Institut Ljubljana (OI LJUBLJANA)
7	Environmental risk factors and breast cancer	Actions undertaken to study the relation between pollutants contamination and cancer risk	INT

All work packages activities are ongoing and connected

1.4. DELIVERABLES AND MILESTONES IN THE FIRST 18 MONTHS

Code*	Wp	Content specification	month
M1.1	1	First Steering Committee meeting	M1-M2
MD.3	2	A leaflet to promote the project must be produced at the beginning	M3
MD.5	2	Each project must have a dedicated web-site / web-pages. This can have a public part and another one accessible only to the applicants.	Since M3
D4.1	4	Survey devoted to collect information to perform spatial analysis, and deprivation index estimate	M4
M2.1	2	Stakeholder mapping exercise for dissemination	M5
D4.2	4	Protocol describing data collection from CRs for breast cancer data	M5
MD.6	1	Consortium agreement (as mandated in the guide for applicants)	M6
M5.1	5	Creation of WP-5 group of experts on SES	M7
M6.1	6	Creation of WP-6 group of experts on spatial analysis	M7
M4.2	4	Complete list of participating CRs	M8
M4.1	4	INT Ethical Committee (M5-M6) and other Ethical Committees approval	M10
M2.2	2	Cooperation agreement of cancer leagues	M15
M4.3	4	Available data from CRs	M15
D5.1	5	Report indicating the available data for estimating deprivation indexes in the various participating countries	M15
D7.1	7	Report on literature review on environmental factors and breast cancer	M15

* MD: Mandatory deliverable; D: Deliverable; M: Milestone

All deliverables and milestone were covered, for the majority, in the month forecast

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

1.5. GANTT

We here copy the GANTT included in the project application:

WP	Task	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36	M37	M38
1	Consortium agreement	MD.6																																					
1	M1.1: Steering Committee meeting																																						
2	Leaflet	MD.3																																					
4	CR survey	D4.1																																					
2	Web site			MD.5																																			
4	Protocol on CR data collection			D4.2																																			
5	M5.1: group of experts on SES																																						
6	M6.1: group of experts on spatial analysis																																						
2	M2.1: stakeholder mapping exercise																																						
4	M4.1: Ethical Committee approvals																																						
4	M4.2: List of participating CRs																																						
4	M4.3: CR data collection																																						
5	Report on deprivation index available data										D5.1																												
6	Report on methods for spatial analysis																																						
7	Report on pollutants and breast cancer										D7.1																												
2	M2.2: agreement of cancer leagues																																						
5	M5.2: deprivation indexes data collection																																						
4	M4.4: database for spatial analysis																																						
7	Report on environmental data available										D7.2																												
2	M2.3: First version of course material																																						
4	Report on data management model														D4.3																								
1	Interim report																		MD.1																				
6	Practical manual for CR's personnel																					D6.2																
1	M1.2: Steering Committee meeting																																						
2	M2.4: alpha & beta testing for courses																																						
7	Protocol for pilot environmental study																								D7.3														
7	M7.1: Ethical Committee approvals																																						
2	Awareness online courses																												D2.1										
4	M4.5: pilot environmental study database																																						
5	Report on European Deprivation Index																												D5.2										
6	Report on breast cancer spatial analysis																												D6.3										
3	Evaluation financial report																																D3.1						
7	Report on pilot environmental study results																																			D7.4			
1-2	Final reports																																					MD.2 & MD.4	

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

2. ACTIVITIES IN THE FIRST 18 MONTHS

2.1. MAIN RESULTS

We followed the GANTT included in chapter 1.5. In particular, the project's major results are described below:

- WP-1: We held the Steering Commitee Meeting (at M2) and completed the Consortiuum Agreement with a 10-month delay (M17)
- WP-2: We realized and distributed the project leaflet (M3), we launched the logo and website (M2), we identified participating Cancer Leagues (M15) and closed an IT contract for the online courses.
- WP-3: 6-monthly reports were released at M6 and M12
- WP-4: the CR survey was made available to all registries at M4, the data collection protocol was made available online at M2, EC approval at INT was obtained at M5, the list of participating CR was defined at M8, data collection started at M7. The majority of CRs collected data and sent them. They were checked and organized in dataset linked to the geographical information (the smallest territorial unit of residence for every case with a sufficiently-defined geo-codification)and to the already available socio-economic information (the European Deprivation Index in its 2001 and/or 2011 version, where available).
- WP-5: the WP experts group was defined at M 7 and the report on deprivation index available data was completed at M15
- WP-6: The WP group was constituted at M7, the creation of a spatial analysis database was started at M10 and for 10 CRs first analysis were performed
- WP-7: the first report on pollutants and breast cancer was completed on time (on the topic of pollutants and BC, at M15). The WP is preparing a second literature review on ecological study methodologies aimed to correlate water and soil pollutants and some health outcomes

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

2.2. *INDICATORS FOR SPECIFIC OBJECTIVES*

Specific Objective Number	SO-1	
Specific Objective	Identify geographical areas in the European Union with higher breast cancer risk (through spatial analysis of CR data)	
Process Indicator(s)	**Target**	**Value at M18**
Number of CRs participating in the CR survey (see methods and means)	>=20	33
Number of CR countries participating in the survey	>=7	9
Output Indicator(s)	**Target**	**Value at M18**
Number of CRs sending data for spatial analysis on breast cancer	>=15	38
Number of CR countries sending data for spatial analysis on breast cancer	>=6	8
Outcome/Impact Indicator(s)	**Target**	**Value at M18**
Number of CRs performing spatial analysis for the first time	>=4	15
Number of CR countries performing spatial analysis for the first time	>=1	2

Specific Objective Number	SO-2	
Specific Objective	Deprivation indexes estimates to be considered in spatial analysis of breast cancer incidence as SES mediator	
Process Indicator(s)	**Target**	**Value at M18**
See SO1 process indicators (CR survey will cover also SES topic)	-	-
Output Indicator(s)	**Target**	**Value at M18**
Number of CRs for which deprivation indexes is used in breast cancer spatial analysis	>=15	33
Number of CR countries for which deprivation index is used in breast cancer spatial analysis	>=6	8
Number of countries for which EDI is used in breast cancer spatial analysis	>=5	6
Outcome/Impact Indicator(s)	**Target**	**Value at M18**
Number of CRs using deprivation indexes for the first time	>=4	17
Number of CR countries using EDI for the first time	>=2	4

Specific Objective Number	SO-3	
Specific Objective	Review the evidence of polluting agents and their impact on breast cancer	
Process Indicator(s)	**Target**	**Value at M18**
Number of published scientific articles considered	>=50	130
Output Indicator(s)	**Target**	**Value at M18**
Number of water/soil pollutants studied in relation with breast cancer	>=15	40
Outcome/Impact Indicator(s)	**Target**	**Value at M18**
Number of reports on literature review	>=1	1

Specific Objective Number	SO-4	
Specific Objective	Environmental study on correlation between soil and water contamination and breast cancer risk	
Process Indicator(s)	**Target**	**Value at M18**
Number of European databases with available data on water/soil pollutants	>=2	NA
Number of national databases with available data on water/soil pollutants	>=3	NA
Number of pollutant indicators available in different countries	>=5	NA
Output Indicator(s)	**Target**	**Value at M18**
Number of pollutant indicators available for pilot environmental study	>=1	NA
Outcome/Impact Indicator(s)	**Target**	**Value at M18**
Number of pilot environmental studies performed	>=1	NA

NA: Not Yet Applicable

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

Specific Objective Number	SO-5	
Specific Objective	Design courses on breast cancer risk factors awareness for young girls/adolescents	
Process Indicator(s)	**Target**	**Value at M18**
Number of target countries where promoting the online course	>=5	5
Number of participants per target country taking part in “alpha test” of online course	>=15	N/A
Number of participants per target country taking part in final “beta test” of online course	>=15	N/A
Mean age of participants to testing phase	14.5 yrs	N/A
Output Indicator(s)	**Target**	**Value at M18**
Number of unique visits to online course per target country	>=5,000	N/A
Engagement rate of Facebook page (or other social media metric)	>=50%	N/A
Mean age of participants to online course	14.5 yrs	N/A
Outcome/Impact Indicator(s)	**Target**	**Value at M18**
Number of unique visitors completing online course per target country	>=4,000	N/A
Number of downloads of additional/complementary information by unique visitor after completion of course per country OR Number of unique visitors following hyperlinks to referral information / partner web pages after completion of course per country	>=1,000	N/A

NA: Not Yet Applicable

2.3. *WORK PACKAGE 1 – COORDINATION*

INT (main partner) co-ordinates the entire project. In particular:

- A Steering Committee (SC) was created to discuss and establish the pillars of the project. Delegates of the associated partner units were invited to be SC members
- The first SC meeting was organized at 19th February 2018. Minutes are in **Annex 1**. All the presentations are available at the website
- The consortium agreement (CA) was signed in June 2019, in delay in respect to the time plan. First page of the CA is in **Annex 2**. The complete CA is in file attached to the report
- Project information is circulated through the network (associated partners, collaborating partners) by means of periodic email newsletters and 6-monthly report (see WP-3)
- The main documents are made available on the website (see WP-2)
- Deliverables and milestone produced by the various WPs are reviewed centrally at INT

	Website address
Work package 1	http://www.wasabysite.it/wp1.html
M1.1 - First Steering Committee	http://www.wasabysite.it/meeting_01.html
MD.6 - Consortium agreement	Not online

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

2.4. *WORK PACKAGE 2 – DISSEMINATION*

ECL coordinates the dissemination activities. In particular:

- INT prepared the website http://www.wasabysite.it, constantly updated by WP-1
- ECL prepared the leaflet in **Annex 3**
- ECL work with all partners to identify existing communication tools and assets. WP lead conducts stakeholder mapping (in **Annex 4**) for purposes of effective and targeted dissemination of the project including synergy with the ENCR communication tools, and the European Parliament, where relevant.
- ECL provided background information and oral presentation on the WASABY project at the Annual General Assembly in September 2017 (prior to project launch). The purpose was to generate interest amongst the national cancer leagues prior to the project launch and gather support for the dissemination actions. A further update was provided to leagues at the General Assembly in November 2018.
- ECL launched an online contest to develop a logo for the WASABY project, which was concluded in M3. The chosen designer of the logo also produced a brand identity kit to be used in the dissemination actions of the project. A tailored PowerPoint presentation slide deck was also produced for use by project partners in their dissemination activities.
- ECL has worked to engage cancer leagues with the WASABY project dissemination and primary prevention action by constituting a small working group (**Annex 5**). A dedicated overview booklet for cancer leagues was produced and disseminated to all 26 members of the ECL network. In total, 10 leagues expressed an initial interest in joining the working group to help guide the content generation of the primary prevention course. BY M18, 5 leagues have confirmed their participation in the group and a further 4 leagues will confirm their participation or not by the time of the first teleconference, in M20. A first mapping of the content development has begun and a first will be discussed at the teleconference scheduled for M20.
- To develop the online primary prevention course for young women and technical partner with experience in the e-health and m-health field has been contracted to provide the support and infrastructure. The contracted partner has an existing application which can be adapted to suit the needs of the WASABY project. This contracts includes a three-year licence period following the conclusion of the project for widespread dissemination and sustainability of the prevention course.

	Website address
Work package 2	http://www.wasabysite.it/wp2.html
MD.3 - Leaflet	http://www.wasabysite.it/material/MD3_WASABY_Leaflet.pdf
M2.1 – Stakeholder mapping exercise for dissemination	http://www.wasabysite.it/material/M2.1_WASABY_Stakeholder_Mapping.pdf
M2.2 – Cooperation agreement of cancer leagues (or other partners)	http://www.wasabysite.it/material/M2.2_WASABY_Cancer_Leagues_Working_Group.pdf

List of meeting abstract

Title	Website address	Location	Date
WASABY - Water And Soil contamination and Awareness on Breast cancer risk in Young Women	http://www.wasabysite.it/meeting_02.html	INT, Milan (Italy)	21.03.2018
Project WASABY: WAter and Soil contamination and Awareness on Breast cancer risk in Young women	http://www.wasabysite.it/meeting_03.html	GRELL, Trento (Italy)	17-19.05.2018
Stima dell'incidenza di tumore alla mammella in giovani donne con modelli SARAR. Primi risultati del Progetto WASABY	http://www.wasabysite.it/meeting_04.html	AIRTUM, Trapani (Italy)	27-29.03.2019
Estimation of incidence of breast cancer in young women with SARAR models. First results of the WASABY Project	http://www.wasabysite.it/meeting_05.html	GRELL, Lisbon (Portugal)	29-31.05.2019

2.5. WORK PACKAGE 3 – EVALUATION

UZL coordinates the evaluation activities strictly connected with the main partner.
WP-3 regularly monitors indicators included in chapter 2.2.

WP-3 decided to produce two short report at M6 (**Annex 6**) and M12 (**Annex 7**) in order to synthesize the status of the project and weigh any delays.

Evaluation conclusion at M6 was: "*All WASABY work packages started their planned work and all ongoing scientific activities perfectly reflect the original GANTT. In general, the participation of associated and collaborating partners is very high. The Steering Committee meeting was extremely useful to decide on major strategies. The only reported delay is the Consortium Agreement, for which M6 proved to be too early a deadline in consideration of the required administrative effort. Of note, the delay does not have any impact on the smooth running of the project activities.*"

Evaluation conclusion at M12 was similar: "*All work packages carried on work as planned; the ongoing scientific activities perfectly reflect the original GANTT. In general, the level of involvement of both the associated and the collaborating partners is very high. For just three CRs, we report a slight delay in obtaining the Ethical Committee approvals. We received data from 10 CRs in three countries; data collection and preparation is ongoing in the majority of the remaining CRs. The only reported delay continues to be the Consortium Agreement. Of note, the delay does not have any impact on the smooth running of the project activities.*"

	Website address
Work package 3	http://www.wasabysite.it/wp3.html
Report at M6	http://www.wasabysite.it/material/Wasaby_Report_M6.pdf
Report a t M12	http://www.wasabysite.it/material/Wasaby_Report_M12.pdf

2.6. *WORK PACKAGE 4 – DATA MANAGEMENT*

INT coordinates the data management activities. In particular:

- A survey was sent to European CRs with the aims of a) identifying the CRs interested to participate in the action and b) collecting information on relevant data availability in the CRs. Results are in **Annex 8**
- A data collection protocol, linked with the production breast incidence maps, was prepared (**Annex 9**). The protocol have a double aim: to describe data collection and to prepare the material for the Ethical Committee approval
- The protocols are first sent to the INT Independent Ethical Committee and where necessary also to the local ethical committees according to CR's requirements (**Annex 10**)
- The collected data were managed at INT. The management included: data quality check, geo-coding procedures where needed, linkage with the socio-economic information (the European Deprivation Index), production of a final dataset for each CR. In next pages we list info on the management of the data for each CR.

	Website address
Work package 4	http://www.wasabysite.it/wp4.html
D4.1 – CR survey	http://www.wasabysite.it/material/D4.1_WASABY_CR_Survey.pdf
D4.2 – Protocol for CR data collection	http://www.wasabysite.it/material/D4.1_WASABY_Data_collection_protocol.pdf
M4.1 – INT Ethical Committee and other Ethical Committees approval	http://www.wasabysite.it/material/M4.1_WASABY_WP4_Ethical_Committees.pdf
M4.2 – Complete list of participating CRs	http://www.wasabysite.it/wp4_cr.html
M4.3 – Available data from CRs	

List of adhering cancer registries (CR)

Cancer Registry	Country	Wasaby referent	Status at 31 May 2019
Gironde	France	Marc Colonna for all French CRs	French CRs will not send data to the coordinator center but they will share results and methods of spatial analysis.
Poitou-Charentes			
Loire-Atlantique et Vendée			
Haute Vienne			
Calvados			
Manche			
Somme			
Lille et sa région			
Bas-Rhin			
Haut-Rhin			
Doubs et Territoire de Belfort			
Côte-d'Or			
Isère			
Hérault			
Tarn			
Guadeloupe			
Martinique			
Schleswig-Holstein	Germany	Ron Pritzkuleit	Schleswig-Holstein CRs will not send data to the coordinator center but they will share results and methods of spatial analysis.
Alto Adige	Italy	Guido Mazzoleni	After local Ethical Committee approval, the geo-coding process is on-going.
Napoli 3 South	Italy	Mario Fusco & Francesca Maria Vitale	Sent all data.
Palermo	Italy	Rosanna Cusimano &Walter Mazzucco	Sent all data.
Parma	Italy	Paolo Sgargi	Sent all data.
Ragusa	Italy	Rosario Tumino	Sent all data.
Siracusa	Italy	Francesco Tisano	Sent all data.
Trento	Italy	Silvano Piffer	After local Ethical Committee approval, the geo-coding process is on-going.
Trapani	Italy	Giuseppa Candela & Tiziana Scuderi	Sent all data. Geo-coding process is on-going.
Umbria	Italy	Fabrizio Stracci	Sent all data.
Varese	Italy	Giovanna Tagliabue & Paolo Contiero	Sent all data
Lithuania	Lithuania	Ieva Vincerževskien	Sent breast cancer data, population data and shape files. Waiting for deprivation index data
Polish National Cancer Registry	Poland	Krzysztof Czaderny	Polish CRs will not send data to the coordinator center but they will share results and methods of spatial

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

Cancer Registry	Country	Wasaby referent	Status at 31 May 2019
Greater Poland	Poland	Maciej Trojanowski	analysis. They are coordinated by the Polish National Cancer Registry, who is a collaborator of the WASABY Project. Ethical Committee approval obtained.
Kielce	Poland	Ryszard Mezyk	
Silesia	Poland	Marcin Motnyk	
Masovia	Poland	Urszula Sulkowska	
Podkarpackie	Poland	Monika Lampart	
Central Portugal	Portugal	Joana Bastos	Sent all data.
Northern Portugal	Portugal	Luis Antunes & Maria José Bento	Sent all data.
Slovenia	Slovenia	Vesna Zadnik	Slovenia CR will not send data to the coordinator center but they will share results and methods of spatial analysis.
Basque Country	Spain	Nerea Larrañaga	Sent all data.
Castellon-Valencia	Spain	Paloma Botella	Sent breast cancer data, population data and shape files. Waiting for deprivation index data.
Girona	Spain	Rafael Marcos-Gragera	Sent breast cancer data, population data and shape files. Waiting for deprivation index data.
Granada	Spain	Maria José Sanchez Perez	Sent all data.
Murcia	Spain	Maria Dolores Chirlaque	Data collection on-going.
Navarra	Spain	Marcela Guevara	Sent breast cancer data, population data and shape files. Waiting for deprivation index data.
Northern Ireland	UK	Anna T. Gavin & David Donnelly	Sent population and shape files. Waiting for data transfer agreement signatures for breast cancer data.

List of cancer registries sending data on breast cancer cases to the coordinating center

Cancer Registry	Country	Period	Incidence period	Nr of cases Age:0-49
Varese	Italy	May 2018	1996-2012	2679
Ragusa	Italy	Sep 2018	2001-2012	515
Umbria	Italy	Oct 2018	2001-2013	2008
Siracusa	Italy	Oct 2018	2004-2013	587
Parma	Italy	Nov 2018	2001-2010	424
Napoli 3 South	Italy	Jan 2019	2008-2015	1514
Palermo	Italy	Mar 2019	2005-2014	1914
Alto Adige	Italy	Mar 2019	2004-2013	747
Trapani	Italy	Apr 2019	2002-2011	575
Lithuania	Italy	Apr 2019	2001-2012	3797
Northern Portugal	Portugal	Nov 2018	2003-2012	5561
Central Portugal	Portugal	Apr 2019	2001-2011	3715
Basque Country	Spain	Sep 2018	2005-2014	3512
Granada	Spain	Sep 2018	2004-2013	1245
Navarra	Spain	Nov 2018	2004-2013	1057
Girona	Spain	Jan 2019	2001-2010	993
Castellon-Valencia	Spain	Feb 2019	2004-2014	936

List of cancer registries collecting data (analysis will not be performed by the coordinating center)

Cancer Registry	Country	Period	Status	Incidence period	Nr of cases Age:0-49
Gironde	France	Mar 2019	Data collected	2008-2015	2308
Poitou-Charentes	France	Mar 2019	Data collected	2008-2015	2513
Loire-Atlantique et Vendée	France	Mar 2019	Data collected	2006-2015	3621
Haute Vienne	France	Mar 2019	Data collected	2009-2015	477
Calvados	France	Mar 2019	Data collected	2006-2015	1285
Manche	France	Mar 2019	Data collected	2006-2015	831
Somme	France	Mar 2019	Data collected	2006-2015	1018
Lille et sa région	France	Mar 2019	Data collected	2008-2015	1286
Bas-Rhin	France	Mar 2019	Data collected	2006-2013	1654
Haut-Rhin	France	Mar 2019	Data collected	2006-2015	1253
Doubs et Territoire de Belfort	France	Mar 2019	Data collected	2006-2015	840
Cote d'Or	France	Mar 2019	Data collected	2006-2015	891
Isère	France	Mar 2019	Data collected	2006-2015	2223
Hérault	France	Mar 2019	Data collected	2006-2015	2075
Tarn	France	Mar 2019	Data collected	2006-2015	672
Guadeloupe	France	Mar 2019	Data collected	2006-2015	n.a.
Martinique	France	Mar 2019	Data collected	2006-2015	n.a.
Schleswig-Holstein	Germany	Apr 2019	Data collected	2001-2015	7513
Greater Poland	Poland	Apr 2019	Data collected	2007-2016	3358
Kielce	Poland	Apr 2019	Data collected	2007-2016	915
Silesia	Poland	Apr 2019	Data collected	2007-2016	3651
Masovia	Poland	Apr 2019	Data collected	2007-2016	4323
Podkarpackie	Poland	Apr 2019	Data collected	2007-2016	1392
Slovenia	Slovenia	Apr 2019	Data collected	2006-2015	2464

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

2.7. *WORK PACKAGE 5 – DEPRIVATION INDEXES*

UNICAEN coordinates WP-5 activities. In particular:

- For the collection of data on deprivation indexes or the deprivation indexes estimate, a list of experts is identified covering each country involved (**Annex 11**)
- The main aim is to identify those participating countries for which a national version of the European Deprivation Index (EDI) is not available. For these countries, we are collecting information on census data availability, according to the data level, in order to estimate the possibility of a deprivation index construction. If EDI is not estimable, other deprivation indexes should be identified and collected.
- At this stage, the indicators for estimating the European Deprivation index in WASABY described in **Annex 12** are:
    - Five countries for which we are sure about the EDI 2011 estimation: France, Spain, Italy, Portugal, Slovenia
    - Three countries for which we envisage the EDI 2011 estimation for the first time. It will be possible according to the availabilities of access to aggregated census data at the pertinent level: Lithuania, Poland, Northern Ireland
    - One country for which we will not be able to do EDI estimation: Germany

	Website address
Work package 5	http://www.wasabysite.it/wp5.html
M5.1 – Creation of WP5 group of experts on SES	http://www.wasabysite.it/material/M5.1_WASABY_WP5_Group_of_experts.pdf
D5.1 – Report on collection of information of already available deprivation indexes and on data necessary to estimate European Deprivation Indexes	http://www.wasabysite.it/material/D5.1_WASABY_Deprivation_Index_available_data.pdf

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

2.8. *WORK PACKAGE 6 – METHODS AND ANALYSIS*

OI LJUBLJANA is responsible for evaluation of methods with a group of experts (**Annex 13**) while INT performs analysis on CR data received.

The first round of analyses were performed at INT, using spatial auto-regressive models with the effect of auto-regressive disorder and exogenous covariates (SARAR models). These models were chosen as the first exploratory analysis model, because they allow to consider covariates in the spatial autoregression. The analyses were performed by the SPREG module in STATA 14.

List of analysis performed with cancer registries data sent to the coordinating center

Cancer Registry	Country	Period	Status
Varese	Italy	May 2018	First analysis sent to registry (**Annex 14**)
Ragusa	Italy	Sep 2018	First analysis sent to registry (**Annex 15**)
Umbria	Italy	Oct 2018	On-going analysis
Siracusa	Italy	Oct 2018	First analysis sent to registry (**Annex 16**)
Parma	Italy	Nov 2018	First analysis sent to registry (**Annex 17**)
Napoli 3 South	Italy	Jan 2019	First analysis sent to registry (**Annex 18**)
Palermo	Italy	Mar 2019	On-going analysis
Alto Adige	Italy	Mar 2019	On-going analysis
Trapani	Italy	Apr 2019	On-going analysis
Lithuania	Italy	Apr 2019	On-going analysis
Northern Portugal	Portugal	Nov 2018	First analysis sent to registry (**Annex 19**)
Central Portugal	Portugal	Apr 2019	First analysis sent to registry (**Annex 20**)
Basque Country	Spain	Sep 2018	First analysis sent to registry (**Annex 21**)
Granada	Spain	Sep 2018	First analysis sent to registry (**Annex 22**)
Navarra	Spain	Nov 2018	On-going analysis
Girona	Spain	Jan 2019	On-going analysis
Castellon-Valencia	Spain	Feb 2019	On-going analysis

	Website address
Work package 6	http://www.wasabysite.it/wp6.html
M6.1 – Creation of WP6 group of experts on spatial analysis	http://www.wasabysite.it/material/M6.1_WASABY_WP6_Group_of_experts.pdf
Status of CR data analysis	http://www.wasabysite.it/wp4_cr.html

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

2.9. *WORK PACKAGE 7 – ENVIRONMENTAL RISK FACTORS AND BREAST CANCER*

INT coordinates WP-7 activities. In particular:

- Scientific literature review on breast cancer and environmental factors was performed (**Annex 23**). Synthesis of literature review is in the following table

Contaminants	number of scientific articles	NO ASSOCIATION scientific articles	POSITIVE scientific articles
PCBs (209 congeners)	38	20	18
DDT, DDD, DDE and principal Organachlorines Compounds	40	33	7
Dioxins (TCDD)	13	8	5
PAHs	14	8	6
PFAA (PFOS, PFOSA and PFOA)	6	3	3
Triazine	8	7	1
Cadmium (Heavy Metal)	6	3	3
Trihalomethanes (THMs)	5	2	3
TOTAL Scientific Articles	130	84	46

- A second literature review literature review was performed to identify various environmental study methodologies applied across the world to evaluate correlation between water & soil pollutants (ex: arsenic in water) and a health outcome (ex: cancer incidence) with data available independently from the aim of the environmental study. Search term research defined in PubMed identified 694 articles. During the first phase of the abstract revision, 122 articles were identified by at least one of three reviewers. After the second revision of entire articles by four reviewers, 40 articles were identified as useful for the WASABY objectives. A scientific article is in preparation
- Revision of online database started. The main databases identified at this stage are: Waterbase Ground Water, E-PRTR and Geochemical Atlas of Europe.

	Website address
Work package 7	http://www.wasabysite.it/wp7.html
D7.1 – Report on literature review on environmental factors and breast cancer	http://www.wasabysite.it/material/D7.1_WASABY_Literature_Review.pdf

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

3. ACTIVITIES IN THE SECOND 18 MONTHS

3.1. NEXT ACTIVITIES BY WORK PACKAGE

We will follow the GANTT included in chapter 1.5.

In particular next activities will be aimed to meet the following aims:

- WP-1: the present interim financial and scientific report submitted at M18 will function as guide for the following 18 months. A second steering committee meeting will be organized in order to asses the main results of spatial analysis and to define the area for the WP-7 pilot study
- WP-2: will focus on the preparation of course material and online courses. WASABY will continue to disseminate results in national and European conferences
- WP-3: will continue to evaluate the project with 6-monthly reports. In the last 6 months of the project WP-3 will perform a detailed evaluation on the financial aspects of the project
- WP-4, WP-5 and WP-6: activities will focus on
  - Comparison between methods to perform spatial analysis. Mainly, SARAR method (used in this phase of the project) will be compared with Bayesian methods, such as INLA, performed in R environment
  - Conclusion of European Deprivation Indexes estimates
  - Conclusion of analysis in various CRs. Meetings or online conferences will be organized with those CRs performing analysis itself (French and Polish CRs)
  - Preparing a report on practical issues to follow for spatial analysis to be performed by cancer registry operators
- WP-7: activities will focus on
  - Preparation of a scientific article on literature review performed in this first phase of the project
  - Evaluation of other available water and soil pollutants databases to be used for ecological studies with breast cancer incidence data analysed by other WPs
  - Definition of a environmental pilot study to be performed in one CR area.

3.2. CONTINGENCY PLAN

We here copy the part of the contingency plan (indicated in the proposal) connected with this forst 18 months of the project, including a column with M1-M18 results and any relevant action that was taken.

Identified Risk	**Likelihood**	**Impact**	**M1-M18 status and activities**
General risks			
Divergent goals/vision	Low	Medium	During the SC meeting we discussed the overall WASABY vision and all members agreed on the steps to be followed within the project
SO-1: Investigate geographical areas in the European Union for higher breast cancer risk (through spatial analysis of CR data)			
The number of CRs may be too low A few CRs declaring interest to participate may not provide data	Low	Low	• The number of participating CRs is satisfactory • 2 CRs (Belgium and Bremen) declaring interest to participate, after 12 months decided not to participate for internal organizational problems
INT and/or CR Ethical Committees does not approve the use of data	Low	Medium	• Some delays of obtaining EC approvals was connected to the new GDPR regulation implementation • In one case (Northern Ireland CR a slight delay is caused by the fact that a Data Transfer Agreement needs to be formally prepared and signed
Change of the licensing of the software used in WP-4 and WP-6	Low	Low	We encountered no problems so far. We focus our activities using open source software (QGis, R). However, we already have a license for ArcGis (instead of QGis) and Stata (instead of R)
No data on census block boundaries are available in some CRs	Medium	Medium	In Germany, spatial analysis on breast cancer incidence would be performed with other (larger) sub-areas such as municipalities and districts.
No data on screening adherence are available	Medium	Low	Spatial analysis on breast cancer incidence is performed on ages < 50 (not covered by screening programmes in the majority of countries)
SO-2: Deprivation indexes estimates in order to use them as possible SES mediator in spatial analysis of breast cancer incidence			
European deprivation index could not be estimated / could be estimated for different sub-areas	Medium	Medium	We are currently considering the possibility of using other (already available) national deprivation indexes

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

ANNEX 1 – M1.1 – MINUTES OF THE 1ST STEERING COMMITTEE MEETING

WASABY SC MEETING 19.02.2018, INT MILAN
MEETING MINUTES WITH LINKS TO WORKING MATERIALS

On Monday 19.02.2018 the 1st Steering Committee Meeting of the WASABY project was organized, in the premises of the Cancer Institute of Milan, INT. Participants included the 7 WP Leaders, researchers involved in the WP work of the partner institutes and a few auditors and experts. The agenda of the day included a morning session - covering presentations on the Horizontal Work Packages by Camilla Amati of WP1 and Paolo Baili (P.I.), David Ritchie of WP2 and a discussion on the Scientific Work Packages (Roberto Lillini, WP4; Tina Zagar, WP6; Guy Launoy and Elodie Guillaume WP5) – and an afternoon session including presentations from WP7 as well as on the visual identity and communication means to be used in the project, e.g., website, leaflet, newsletters, presence in congresses and internal communications. For Meeting Materials, including agenda, participants, project goals, draft leaflet, GANNT, WP4 questionnaire to CRs and proposed data collection protocol, please see here.

WP1 presentations covered the role of the Steering Committee, the aims of the day, the project's target groups, working structure, goals and deliverables, as well as the timeframe and administrative requirements. To see presentations, please click here: 1, 2.
MAIN MESSAGE: The shared vision of working country by country was agreed. Both the scientific and administrative issues reported in the Grant Agreement are to be endorsed by the single partners, but the WP1 team is available for questions at any time. The next Steering Committee Meeting will be organised at around M 24, the interim report will be at M 18, the final report will be at M 38. A consortium agreement will be sent to all legal offices for signature soon.

For WP2, ECL's background and scope were presented in connection with the WASABY objectives. To see the presentation on WP2, please click here.
MAIN MESSAGE: WP2 will be able to help the partners disseminate year 1 and 2 activities. Specific collaboration on the online courses to be developed in the 3rd year will be organized later on into the project, e.g., in the occasion of the next SC Meeting.

WP4's presentation covered the data management objectives and actions, a state of the art on participating Cancer Registries and survey results, and a discussion on the proposed data collection protocol. To see WP4 presentation, please click here .
WP6's presentation included the description of the Slovenian experience on geo-coding and spatial analysis, and a discussion on WP6 actions and strategies with the participants. To see WP6 presentation, please click here.
WP5 colleagues illustrated the European Deprivation Index and the related activity conducted in Caen with 6 countries, and a working strategy for WP5 in the Wasaby project. To see WP5 presentation, please click here.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

The morning decisions taken are summarised here.

In particular, the CR inclusion criteria were revised as follows:

- work will be initially started with those CRs able to analyse data by "smallest unit" (S.U.) (not by municipality);
- geo-coded residence addresses is intended as residence address at the date of diagnosis;
- as the number of participating CRs, the number of EU countries involved in the project, and the number of CRs performing spatial analyses for the first time correspond to set targets to be reported at project M7 and 10 (milestones 4.2 and 4.4), the possible involvement of further CRs/further countries shall be considered only after M10.

Please note a modified version of the Data Collection Protocol, according to suggestions and observations coming from discussion, is sent to the Steering Committee members with the present document, for final approval. The document will be circulated to the participating Cancer Registries in connection to possible Ethical Approval Applications and to start registry-specific discussion with WP4.

WP7 colleagues illustrated the background for the Environmental Study WP, the main methodological aspects of the search for scientific literature as well as national and international databases on water and soil pollutants. Scientific literature review should be also focused on publications including ecological studies evaluating relationship between water&soil pollutants and cancer risk (not strictly on breast cancer). The strategy for the pilot environmental study will be discussed in the next meetings. To see WP7 presentations, please click here 3, 4, 5 .

Other decisions taken during the afternoon, covering dissemination opportunities and publication policy, were agreed, as illustrated here. These covered the project's website domain, the project logo choice, the presence of a WASABY poster or oral presentation at the GRELL meeting in Trento, Italy (for info, see here), and at the ENCR annual Meeting in Copenhagen, Denmark (for info, see here) a publication policy article to be included in the Consortium agreement.

The project Leaflet was illustrated, comments from all partners are welcome. A graphic designer will be contacted by ECL: a revised proposal will soon be circulated and linked in an updated version of this document.

A WP3 work plan was not addressed in the meeting, because WP3 activity will start at M3 and the first deliverable coincide with the preparation of the Interim Report in collaboration with WP1 (M 18). Nonetheless, a representative of WP3 was present at the meeting and provided useful elements for the discussion on the activities of the various WPs.

NEXT STEPS: from now every WP leader will start the relevant contacts with WASABY participants and external experts in order to structure their scientific activity and meetings as needed. These actions will be copied to WP1 and WP4, in order to reflect the interactive character of the Project.

ANNEX 2 – MD.6 – FIRST PAGE OF THE CONSORTIUM AGREEMENT

WASABY EU 3rd Health Programme-Workplan 2016 - Call ID:PP-2-2016

CONSORTIUM AGREEMENT

for the execution of the action entitled

"Water and soil contamination and Awareness of Breast Cancer risk in Young women"

in short **"WASABY"**

2019-05-21, DESCA2020 CA x Wasaby, FINAL

WASABY

ANNEX 3 – MD.3 – LEAFLET

WHO IS
INVOLVED?

PROJECT PARTNERS:

- National Cancer Institute (Italy) [project lead]
- The Association of European Cancer Leagues (ECL)
- Institute of Oncology Ljubljana (Slovenia)
- University of Caen Normandy (France)
- University of Lübeck (Germany)

+ Plus, a growing number of cancer registries in Europe.

Why not get involved?
Email: wasaby@wasabysite.it

"Depending on when in her lifespan a woman integrates risk-reduction behaviours, the majority of breast cancer can be prevented"

Colditz G. CA Cancer J Clin, 2014

WASABY

Visit our website to learn more:
www.wasabysite.it

EU funding disclaimer:
This project has received funding from the 3rd European Union Health Programme 2014-2020 under Grant Agreement PP-2-5-2016 (#769767)

Co-funded by the Health Programme of the European Union

WASABY

WATER
AND
SOIL CONTAMINATION
AND
AWARENESS
ON
BREAST CANCER RISK
IN
YOUNG WOMEN

WHAT IS
WASABY?

WASABY is a 3-year EU funded project beginning in January 2018.

The project aims to include the design of a model able to identify areas with higher cancer rates, to study whether pollutant contamination may be a cause for increased cancer risk.

THE MAIN OBJECTIVES ARE:

TO DESIGN
procedures for producing cancer maps

TO CONDUCT
ecological studies on association between pollution and increased cancer risk.

WHAT WILL
WASABY DO?

During the first two years, WASABY will produce breast cancer risk maps identifying areas at higher risk in Europe using specific geographic information systems and ad hoc models of spatial analysis.

An environmental study will be piloted on water and soil pollutants and breast cancer risk in the third year.

During the final year, WASABY will develop online primary prevention courses to increase awareness of breast cancer risk amongst girls aged 12 to 19 years old.

ADDED VALUE
OF WASABY

- Cooperation and exchange between participating population-based cancer registries to produce an innovative methodology
- Improve the use of deprivation indexes in spatial analysis and enlarge the utilisation of the European Deprivation Index
- Increase spatial analysis activities across European cancer registries to distinguish clusters of breast cancer risk with a focus on young women
- Pilot the connection between databases of water and soil pollutants and cancer registry databases
- Prepare a useful web instrument to promote breast cancer preventive actions among girls aged 12 to 19 years old

ANNEX 4 – M2.1 – STAKEHOLDER MAPPING EXERCISE FOR DISSEMINATION

Introduction

WASABY (Water and Soil Contamination and Awareness of Breast Cancer Risk in Young Women) is 3-year EU funded project co-financed by the Third EU Health Programme 2014-2020. The project aims to design of a model able to identify areas with higher cancer rates, to study whether pollutant contamination may be a cause for increased cancer risk and develop online primary prevention courses to increase awareness of breast cancer risk amongst girls aged 12 – 19 years old.

The WASABY project partners are:

	Applicant organisation name	**Acronym**	**Country**
1 (Coordinator)	FONDAZIONE IRCCS ISTITUTO NAZIONALE DEI TUMORI	INT	Italy
2	ASSOCIATION EUROPEENNE DES LIGUES CONTRE LE CANCER ASBL	ECL	Belgium
3	UNIVERSITÄT ZU LÜBECK	GER	Germany
4	UNIVERSITE DE CAEN NORMANDIE	FRA	France
5	ONKOLOSKI INSTITUT LJUBLJANA	SLO	Slovenia

The WASABY project has three distinct target groups:

1. **Cancer registries** – key actors and receivers of WASABY project outputs and outcomes.
2. **Public health managers** – public health policy and decision makers in Europe;
3. **Young girls and adolescents** – girls aged 12-19 years old who will be targeted for the online training course.

In order to reach the target groups and to communicate the progress and achievements of the project, it is necessary to identify key stakeholder organisations. A stakeholder is understood to be someone who has an interest in an organisation, programme, or project. Stakeholders either affect, or are affected by, the tasks and outcomes of the organisation, programme, or project. In the context of WASABY, stakeholders have the potential to be broadly defined across three axes:

- micro scale - a key target group that is directly involved in, or engaged by, the project;
- meso scale - intermediaries that are communicated with to reach one or more of the target groups; and
- macro scale - organisations who can help disseminate the existence and achievements of the project.

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

Target groups

Cancer Registries

Over 150 European Cancer registries (CRs) in the EU intercept the main data flows generated by administrative and healthcare facilities to provide cancer basic indicators as incidence, mortality, survival, and prevalence, which constitute a key tool for estimating the burden of cancer in populations. In WASABY, CRs can investigate a topic of high relevance in public health: the study of cancer risk using geographical analysis on European level. Key activities focus on a) analysing CRs data, b) the definition of a replicable model of spatial analysis to be used by CRs and c) on correlating CRs data with environmental data. For these actions, CRs are both data providers and target of results.

Given the crucial role of this group in the project, CRs will be addressed at the micro scale by direct email contact, via professional contact (including at conferences, events, and seminars), and via the newsletters and communication channels of CR networks, such as the European Network for Cancer Registries (ENCR). In terms of stakeholder identification for project communication purposes (see annex for list of organisations), CRs can be categorised as follows:

- CRs participating in the project
- CRs not participating in the project
- Networks of CRs.

Public health managers

Public health managers are understood as those actors directly implicated in the frame of national and European-wide policy-making or shaping public health-relevant legislation at the national and European levels. Public health managers have a significant interest in the WASABY project in two ways: firstly, by aiming to study, standardise and promote a new analytical method among the CRs in Europe, WASABY has the improvement of the cancer information system as one of its expected outcome, which will offer public health managers a key element for the improvement of the European citizens' health. Secondly, with the production of breast cancer maps, national and regional health policy makers will receive a tool enabling them to target breast cancer preventive actions at best which have the potential to translated to any cancer site.

Public health managers will be primarily addressed at the meso scale, meaning that this target group will be communicated to predominantly by intermediary organisations, networks, and groups. Indicative channels for communication will come through engagement with the European Parliament interest group on Cancer – MEPs against Cancer; the European Committee of the Regions; WHO networks such as the Health Cities network, and Regions for Health network.

In terms of stakeholder identification for project communication purposes (see annex for list of organisations), this target group can be reached through the following:

- EU institutions, advisory bodies, and agencies
- Health professionals' associations and networks
- International agencies
- Public sector networks

NGOs in the cancer, public health and environmental domain can also be helpful intermediary organisations to reach this target group.

Young girls & adolescents

One of the main objectives of the project is to design and implement an online training module for school-age girls (aged 12-19 years old) that promotes breast cancer risk prevention actions and awareness. The training will be developed in consultation with a sample of the target group and targeted at specific countries and localities selected based upon the outputs of the project's work with CRs.

This target group will be addressed at both the meso scale and macro scale, meaning that communication will be focused on intermediary organisations, networks, and groups to make specific efforts reach the target age group, and global dissemination of the project to reach possible intermediaries not directly identified by the project, and potentially the target age group themselves.

Indicative channels for communication will come via the engagement of cancer leagues in their health education outreach work, whilst collaboration of youth and education NGOs to disseminate the project through existing communication channels. At the EU level, tools such as the health policy platform and European Youth Portal will offer a forum for reaching stakeholders capable of informing the target groups about the training outputs of the WASABY project.

In terms of stakeholder identification for project communication purposes (see annex for list of organisations), this target group can be reached through the following:

- Cancer specific NGOs (notably cancer leagues)
- Educational & youth specific NGOs
- Public sector networks (for example, schools for health network).

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016

Communication channels and tools

WASABY COMMUNICATION TOOLS & CHANNELS				
METHOD	**TARGET**	**WHY**	**WHERE**	**WHEN**
Brand identity pack Professionally designed project logo, PowerPoint slide deck and other promotional tools (e.g. Facebook banner, business card, etc.) for use for project partners.	Project partners	Promoting project	Project website Email	M3
Project Information Leaflet Short, attractive, including a section adaptable to local sites' organizational features, including information on the trial and reference to the project's website.	Interested stakeholders	Inform about project	Printed and available for dissemination at conferences PDF available on project website and partners' websites	M5
Briefing note on WASABY Concise introduction to the project to inform cancer leagues, as a key stakeholder, about WASABY and encourage their active participation in the project implementation.	Cancer Leagues	Inform about project and encourage participation Engaging stakeholders	Available in PDF on project website Disseminated by email to cancer leagues	M9
Explainer Video "Cartoon" format, simple and attractive, explaining the project	Interested stakeholders Public	Inform about project Raising awareness of topics Engaging stakeholders	Project website	M13
Social Media Promote the courses, in line with the needs and preferences of the target group. The project acronym (WASABY) used to design a clear and distinctive identity for the online courses and project identity that will be communicated to the target group through the most popular social media platforms.	Interested stakeholders Target group of girls and adolescents	Inform about project Encourage interest in online training course	Facebook &/or Instagram	From M13
Lay version of final report Easy guide to the project's main achievements and results.	Public Stakeholder organizations	Project deliverable	Project and partners' website Stakeholder websites	M36

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

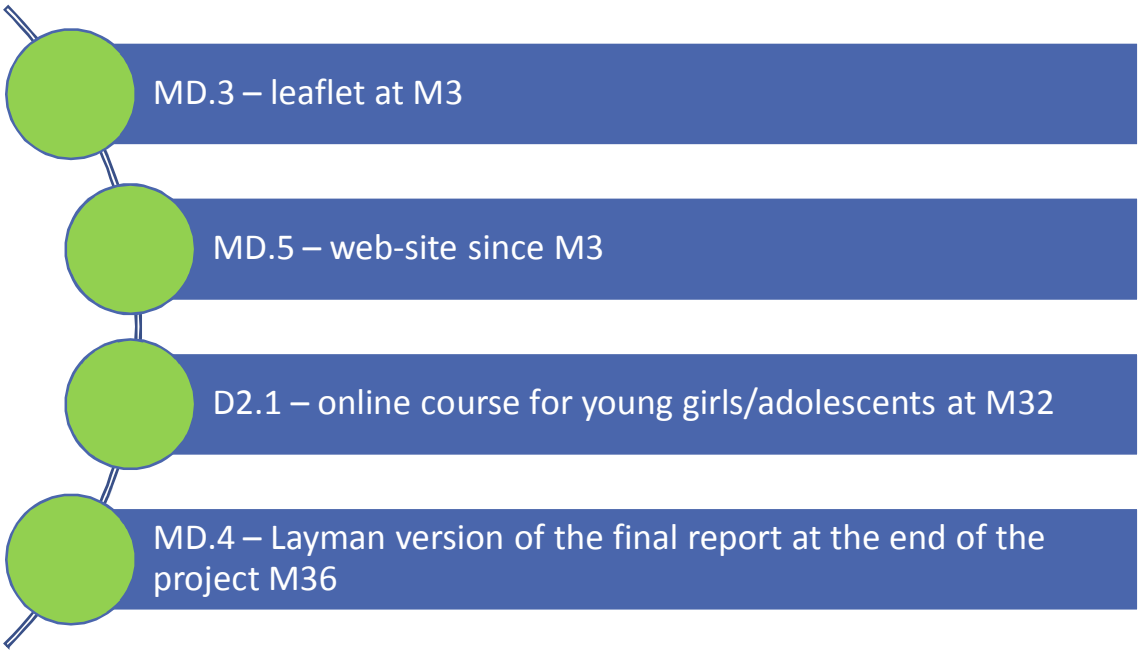
METHOD	TARGET	WHY	WHERE	WHEN
Promotional flyers and posters Tailored materials, infographic leaflets, and posters will be provided in the language of the target country which will aim to refer the target group to the online course.	Interested stakeholders Target group of girls and adolescents	Inform about project Encourage interest in online training course	Printed and available for dissemination at conferences PDF available on project website and partners' websites	M24
Online training course Develop an online primary prevention courses to increase awareness on breast cancer risk to girls from 12 to 19 years of age.	Girls aged 12-19 years old	Project deliverable	Special section of website or dedicated micro-site	M28
Press releases Press releases will be drafted in English and the language[s] of the target countries to promote the key developments in the project.	Media	Inform about project Disseminate results	Project website Conferences E-mail	Results delivery Launch of online courses "Opportunity windows" e.g. breast cancer awareness month, World Cancer Day, European Week Against Cancer, national breast cancer days/weeks, etc.
Newsletter(s) Content regarding the project will be provided for project partner's newsletters and newsletters of stakeholder organizations.	Health Professionals CRs Policy makers Stakeholder organizations	Updating on project progresses Disseminating results and recommendations Putting relevant issues in policy makers' agenda	E-mail Project website Partners' websites and social media accounts	Periodically, along the project's timeline "Opportunity windows"
Scientific articles, letters to editor, commentaries, presentations Key dissemination outputs by project partners engaged in the scientific output of the project	CRs Scientific community Health Professionals Policymakers	Raising awareness Promoting the project Updating on progress Disseminating results and recommendations	Scientific journals Specialized press (grey lit.) Conferences and events	Results delivery "Opportunity windows"

Co-funded by the Health Programme of the European Union

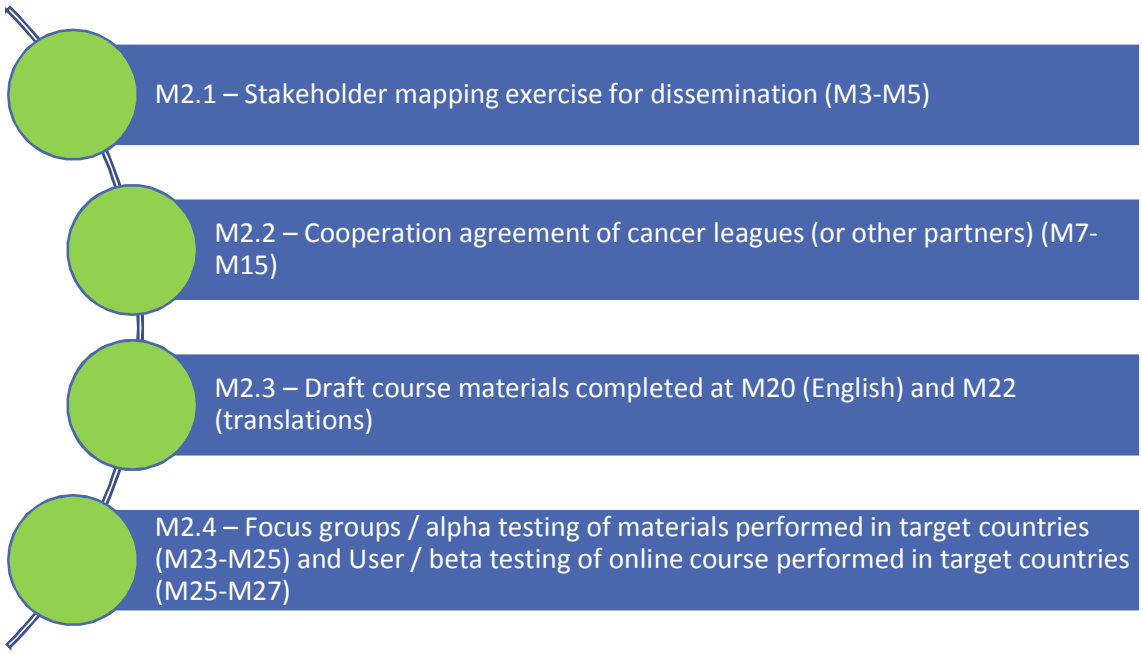
EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Timeline

As per the project protocol, the deliverables linked to the dissemination work package (WP2) of the WASABY project are listed below:



The milestones to be reached by this WP are:



ANNEX 5 – M2.2 – CANCER LEAGUES WORKING GROUP

Creation of a Working Group of Cancer Leagues for discussion on WASABY implementation

This working group will discuss with ECL how to develop the WASABY online courses for adolescents and young women.

Cancer Leagues that are part of the WASABY working group are:

- Belgium - Kom Op Tegen Kanker
- Spain - AECC
- Spain – Catalonia – FECEC
- UK Cancer Focus NI
- France - La Ligue Contre le Cancer
- Cyprus- Pasykaf

A final version will be made available by mid April, possibly including the following groups once they confirm their interest or availability:

- Slovenian Cancer Association (phone call arranged 08.04.2019)
- Portuguese League against Cancer (internal issues are delaying decisions)
- Romanian Cancer Society (v. small so not sure of capacity)
- Polish League against Cancer (v. small so not sure of capacity)

ANNEX 6 – REPORT AT M6

Synthetic report of the Wasaby activity at Month 6
3rd July 2018

Introduction

Wasaby started on 1st January 2018 and has the mandate to produce detailed interim report at Month 18. The principal investigator (P.I.) retains useful to update the partners (associated and collaborating) and the European Commission with this synthetic report underlying how the project is going on in respect of the GANTT included in the application.

GANTT - First 6 months

WP	Task	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
1	Consortium agreement	MD.6									
1	M1.1: Steering Committee meeting										
2	Leaflet	MD.3									
4	CR survey	D4.1									
2	Web site			MD.5							
4	Protocol on CR data collection			D4.2							
5	M5.1: group of experts on SES										
6	M6.1: group of experts on spatial analysis										
2	M2.1: stakeholder mapping exercise										
4	M4.1: Ethical Committee approvals										
4	M4.2: List of participating CRs										
4	M4.3: CR data collection										
7	Report on pollutants and breast cancer							D7.1			
5	Report on deprivation index available data							D5.1			
6	Report on methods for spatial analysis							D6.1			

CR: Cancer Registry

Concluded activities

Type	Code	Due	Title	Description	Final document	Final date	Synthesis
Milestone	M1.1	M2	1st SC Meeting	Steering Committee meeting	www.wasabysite.it/meeting_01.html	19.02.2018	✓
Deliverable	D4.1	M4	CR survey	Questionnaire results on available data in various CRs to produce breast cancer incidence maps	PDF	30.04.2018	✓
Deliverable	D4.2	M5	Protocol for CR data collection	Details of data to be collected for breast cancer incidence maps	PDF	28.02.2018	✓
Milestone	M4.1	M6	INT Ethical Committee	D4.2 was approved by INT ethical committee	-	13.04.2018	✓
Milestone	M2.1	M5	Stakeholder mapping exercise	Identification of stakeholders for the project	PDF	09.07.2018	✓
Deliverable	MD.5	M3	Web-site	Online since 28.2.2018	www.wasabysite.it	28.02.2018	✓
Deliverable	MD.3	M3	Leaflet	First version published online on 28.2.2018	PDF	22.05.2018	✓

INT: Fondazione IRCCS "Istituto Nazionale dei Tumori" (Main Partner)

Ongoing activities

Type	Code	Due	Title	Description	Expected final date	Synthesis
Deliverable	MD.6	M6	Consortium agreement	This activity is delayed. A draft model is under discussion at the INT legal office	Nov-Dec 2018	✖
Milestone	M5.1	M7	Creation of WP-5 group of experts on SES	A list of experts to be invited is already defined	31.07.2018	✔
Milestone	M6.1	M7	Creation of WP-6 group of experts on spatial analysis	A list of experts to be invited is already defined	31.07.2018	✔
Milestone	M4.1	M10	Local Ethical Committees approval	In some cases, CRs are requested to obtain the approval of local Ethical Committees/Privacy Commissions: i.e., Belgium; Bremen (Germany); Alto Adige, Siracusa, Trento and Umbria (Italy); Northern Portugal; Navarra (Spain)	Oct-Dec 2018	✔
Milestone	M4.2	M8	List of participating CRs	Current list is available at: http://www.wasabysite.it/wp4_cr.html	31.08.2018	✔
Milestone	M4.3	M15	CR data collection	Data from Varese CR was submitted and first analysis were performed	31.03.2019	✔
Deliverable	D7.1	M15	Report on pollutants and breast cancer	Literature review is ongoing since January including 62 articles, 21 reviews, 3 IARC Monographs	31.03.2019	✔
Deliverable	D5.1	M15	Report on deprivation index available data	A survey to CRs was performed in order to map and evaluate the geographic level of availability of socio-economic census data	31.03.2019	✔
Deliverable	D6.1	M20	Report on methods for spatial analysis	The analysis on Varese CR data were used to identify a list of methodological aspects to be discussed as part of WP-6 activities	31.08.2019	✔

Dissemination activities

Acitivity	Link
Oral presentations at the Annual AIRTUM and GRELL Meetings 2018	http://www.wasabysite.it/meeting_03.html
Link to the WASABY project page on the ENCR (European Network of Cancer Registries) website	http://www.encr.eu/partners-links
As part of collaboration with the JRC-ENCR, the start of the WASABY Project was advertised on the March 2018 ENCR Newsflash which is circulated to all ENCR members and major cancer registration stakeholders, thus ensuring the involvement of European Cancer Registries in the activities. Updates of progresses will follow	not public
The WASABY Leaflet is regularly made available in wide scientific events involving epidemiologists, cancer registry staff and national health authorities	

Conclusion

All WASABY work packages started their planned work and all ongoing scientific activities perfectly reflect the original GANTT. In general, the participation of associated and collaborating partners is very high. The Steering Committee meeting was extremely useful to decide on major strategies (http://www.wasabysite.it/meetings/01_SC_Milan/Wasaby_SC1_Minutes.pdf). The only reported delay is the Consortium Agreement, for which M6 proved to be too early a deadline in consideration of the required administrative effort. Of note, the delay does not have any impact on the smooth running of the project activities.

ANNEX 7 – REPORT AT M12

Synthetic report of the Wasaby activity at Month 12

Introduction

The WASABY project started on 1st January 2018. Although a report is due at M 18, with the present short report the P.I. wishes to update the partners (associated and collaborating) and the European Commission on the project state of the art, using the GANTT prospect included in the application as a basis.

GANTT - Second semester

WP	Task	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
4	M4.1: Ethical Committee approvals										
4	M4.2: List of participating CRs										
4	M4.3: CR data collection										
4	M4.4: database for spatial analysis										
5	M5.1: group of experts on SES										
5	Report on deprivation index available data			D5.1							
5	M5.2: deprivation indexes data collection										
6	M6.1: group of experts on spatial analysis										
6	Report on methods for spatial analysis			D6.1							
7	Report on pollutants and breast cancer			D7.1							
7	Report on environmental data available				D.7.2						
2	Web site			MD.5							
2	M2.2: Agreement of cancer leagues										
2	M2.3: First version of course material										

CR: Cancer Registry

Overview of completed activities

Type	Code	Due	Title	Notes	Final document	Final date	Synthesis
Milestone	M4.2	M8	List of participating CRs	The final list is online. The Belgian CR had to withdraw	www.wasabysite.it	August 2018	✓
Milestone	M5.1	M7	Creation of WP-5 group of experts on SES	The list of experts is online, WP5 leader will now begin the group discussion	PDF	31-08-2018	✓
Milestone	M6.1	M7	Creation of WP-6 group of experts on spatial analysis	The list of experts is online, WP6 leader will now begin the group discussion	PDF	18-07-2018	✓

CR: Cancer Registry

Ongoing activities

Type	Code	Due	Title	Description	Expected final date	Synthesis
Deliverable	MD.6	M6	Consortium agreement	This activity is delayed. A draft model is under discussion at the INT legal office	Apr-May 2019	✖
Milestone	M4.1	M10	Local Ethical Committees approval	The majority of CRs have obtained the Ethical Committee approvals. Delays are reported for: Bremen (D), Trento (I), and Greater Poland Registry (PL).	Mar 2019	PDF
Milestone	M4.3	M15	CR data collection	10 CRs have sent their data i.e. Italy: 5; Spain: 4; Portugal: 1. Draft analyses were performed on data from 6 CRs. Regular contacts are ongoing between WP4 leader and all remaining CRs. http://wasabysite.it/wp4_cr.html	31.03.2019	✔
Milestone	M4.4	M28	Database for spatial analysis		31.03.2020	
Deliverable	D7.1	M15	Report on pollutants and breast cancer	WP7 group concluded a literature review. http://wasabysite.it/wp7_list.html A draft report was prepared.	31.03.2019	✔
Deliverable	D7.2	M22	Report on environmental data available	WP7 initiated a literature review on ecological methods for studying the relationship between pollutants and epidemiological indicators such as incidence and mortality. 650 articles were identified.[1]	31.10.2019	✔
Deliverable	D5.1	M15	Report on deprivation index available data	A survey was sent to all partners to determine the smallest geographical unit for which census data are available and the access condition. It also includes information about indexes already developed for all countries. On nine countries, eight have answered.	31.03.2019	✔
Milestone	M5.2	M28	Deprivation indexes data collection	EDI-2011 was updated for France and Portugal, calculated for Slovenia and is presently ongoing for Italy and Spain. Contacts for EDI estimate in Northern Ireland began in Dec 2018. It will be developed for other countries after the previous one.	31.03.2020	✔
Deliverable	D6.1	M20	Methods for spatial analysis	WP5 and WP6 leaders met in November, in an initial discussion on methods took place.	31.08.2019	✔
Milestone	M2.2	M15	Agreement of cancer leagues	In progress: several cancer leagues have expressed their interest. Confirmation is expected Mid-February 2019	31.03.2019	✔
Milestone	M2.3	M22	First version of course material	Technical providers are being contacted for price quotations with a decision expected end of January Mid-February 2019. Mapping of comparable actions underway	31.10.2019	✔

CR: Cancer Registry; EDI: European Deprivation Index [1] The defined literature review criteria is: (spatial analysis OR geographic analysis OR GIS) AND (water pollution OR water pollutants OR soil pollutants) AND (cancer registry OR population-based OR estimate OR estimating OR cancer incidence OR cancer mortality)

Conclusion

All work packages carried on work as planned; the ongoing scientific activities perfectly reflect the original GANTT. In general, the level of involvement of both the associated and the collaborating partners is very high. For just three CRs, we report a slight delay in obtaining the Ethical Committee approvals. We received data from 10 CRs in three countries; data collection and preparation is ongoing in the majority of the remaining CRs. The only reported delay continues to be the Consortium Agreement. Of note, the delay does not have any impact on the smooth running of the project activities.

ANNEX 8 – D4.1 – CANCER REGISTRY SURVEY RESULTS

1. Background and aims

In the framework of the WASABY Project, the first operative step was devoted to identify the Cancer Registries (CRs) willing to participate to the study. To this aim, the following information was considered:

- Patients data availability: covered area, time range, geo-coding of the information and most disaggregated geographical level by which data were available, cost for organizing and sharing them in the format requested by WASABY;
- Population data availability: as above;
- Availability of socio-economic information and, specifically, deprivation indices (local, national, etc.): as above;
- Maps availability and any cost to be sustained to obtain them;
- Declared interest in participating in the Project;
- Previous experiences in studies connecting cancer and environmental issues with known potential pollution sources in the area covered by every CR;
- Administrative authorities (local and national) of reference for the CRs.

The above was asked via a preliminary questionnaire to 100 European CRs belonging to the European Network of Cancer Registries (ENCR) between the end of 2016 and the start of 2017. 30 CRs from 9 countries declared their interest to participate (Tab. 1).

Tab. 1 – First round (2016-2017) of CRs interested in participating to WASABY.

Nation	Cancer Registry	Nation	Cancer Registry
Belgium	Belgium	Lithuania	Lithuania
Germany	Bremen	Poland	Greater Poland
	Hamburg		Kracow
	Munich		Kielce
	Schleswig-Holstein		Lower Silesia
Italy	Firenze-Prato		Silesia
	Friuli Venezia-Giulia	Portugal	Central Portugal
	Napoli 3 South		Northern Portugal
	Palermo	Slovenia	Slovenia
	Parma	Spain	Basque Country
	Ragusa		Castellon-Valencia
	Siracusa		Girona
	Trento		Granada
	Umbria		Murcia
	Varese	UK	Northern Ireland

The number of CRs declaring their willingness to participate satisfied the number limits (nations and CRs) expressed as objectives in the project proposal presented to the EU Commission.
After the EU Commission approved the funding of the WASABY Project, a slightly revised version of the questionnaire was re-submitted to the 30 CRs who were interested in being involved in the study.

This new submission had two aims:

- To better define the sub-areas covered by European CRs for which it would be possible to collect geo-coded residence address (coordinates or small areas such as postal codes). This was necessary to assign to each case the smallest possible area on which the risk would be estimated and to identify the risk areas in as much detail as possible accounting for population characteristics and the socio economic status. Moreover, the decision of sub-areas also depended on the availability or potential availability of a national deprivation index.
- To update the information of the CRs, particularly regarding any changes in CR's staff for the specific project.

After the first WASABY Steering Committee (held in Milan on the 19th of February 2018), five more questions were added to the questionnaire and separately sent to the participating CRs, who already answered to the re-submission.
These five questions mainly dealt with issues regarding privacy and Ethical Committee permissions, due to the single CRs, because the discussion during the Steering Committee remarked that CRs in the same nations could be tied to provide different kinds of permission, in order to share data and analysis for the study.

Therefore, the CRs Survey questionnaire reached its final version, presented in the following paragraph.

WASABY

2. The Survey questionnaire

FEASIBILITY SURVEY for the DG SANTE PROJECT PP-2-5-2016 WASABY

Water **A**nd **S**oil contamination and **A**wareness on **B**reast cancer risk in **Y**oung women

IMPORTANT NOTE: THIS IS NOT A DATA REQUEST SURVEY

The survey aims to explore interest and possibility of CR in contributing to the study.CR data handling will be regulated by the responsible ethical committees and in conformity with the EC General Data Protection Regulation (2016/679)

REGISTRY (CR) _____ **SURVEY COMPILED BY** _____ **DATE** _____

1. In the past, has your CR been involved in studies on environmental pollution and cancer (ex: ca-co, spatial analysis, etc)? Would you be interested to become involved in one study on this topic?

   ☐ Yes, and we would be interested ☐ Never, but would be interested ☐ No, and not interested

2. Please indicate below what is the minimum administrative area (from here on, "**AREA"**) on which the CR can estimate incidence rates:

   ☐ zip code ☐ census block ☐ other
   ☐ municipalities ☐ provinces *please specify* _____ _____

3. Please indicate below what is available in your database at date of diagnosis for each incident case:

Residence address	☐	Yes	☐	Only for selected groups (request for specific studies)	☐	No
AREA of residence	☐	Yes	☐	Only for selected groups (request for specific studies)	☐	No
Geocode of residence	☐	Yes	☐	Only for selected groups (request for specific studies)	☐	No

4. If none of the above applies, please let us know whether you would be able to collect residence address at date of diagnosis from the General Register Office/ Vital Statistic Office:

   ☐ Yes it is possible ☐ No it is not possible ☐ Our CR is not interested

5. What is the name of the official body that defines the AREA in your country? _____
6. Do you know which is the average number of inhabitants living in the AREA? _____
7. If you were to perform a geo-coding of addresses or analysis by AREA, which would be the body able to link resident address and AREA?

   ☐ The CR can perform the link via its ordinary sources and/or software available
   ☐ I do not know at the moment, but if the project starts I shall acquire the information
   ☐ The CR can perform the link by paying external sources and/or software
   ☐ This link is a free of charge activity performed by _____
   ☐ This link is performed by paying the activity of _____(If possible, please provide a cost estimate:_____)

8. Please indicate below the answer relevant to your CR:

   ☐ data on AREA population are already available at CR per each calendar year
   ☐ data on AREA population are already available at CR per sporadic calendar years (ex: census year)
   ☐ data on AREA population can be requested free of charge to _____
   ☐ data on AREA population are purchasable from _____
   Do you have idea of the costs? _____
   ☐ I do not know, but if the project starts I will acquire the information

9. To your knowledge, is there a deprivation index available for your country?

   ☐ No
   ☐ I do not know, but if the project starts I will acquire the information
   ☐ Yes, it is available at census block level but at the moment our CR does not have in its databases
   ☐ Yes, it is available at municipality level but at the moment our CR does not have it into its databases
   ☐ Yes, the CR has a deprivation index at municipality/census block level

10. Please indicate below the answer relevant to your CR

- ☐ Maps with AREA boundaries geo-codes are already available at CR
- ☐ I do not know, but if the project starts I shall acquire the information
- ☐ Maps with AREA boundaries geo-codes can be obtained free of charge from _____
- ☐ Maps with AREA boundaries geo-codes are purchasable from _____

Do you have idea of the costs? _____

11. Are there any chemical industries, military bases, or other recognized sources of water/soil pollution present in the area covered by your CR?

- ☐ None to my knowledge
- ☐ Yes, but correlation with cancer was never studied
- ☐ Yes, correlation with cancer is/it has been under study, yet never with breast cancer in young women
- ☐ Yes, correlation with cancer is/ it has been under study (including breast cancer in young women)

12. Please report the calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU). _____
13. Please report the calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection). _____
14. Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally. _____ _____ _____
15. Detail limits in publishing maps (see confidentiality above). _____ _____ _____
16. Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?_____

MANY THANKS FOR YOUR TIME AND COLLABORATION!

3. The results of the final Survey questionnaire: the Cancer Registries who confirmed their participation

After the re-submission of the questionnaire, not all the CRs confirmed their participation in the study.
Most of the reasons for this were due to changed conditions in the operating practice, transformations of the CR's identity (i.e., extension of the covered area), changes in privacy rules governing the data privacy and sharing.

The final list of participating CRs is reported below, with a synthesis of available data and corresponding lowest geographical level they can provide (Tab. 2a & 2b).

Twenty four (out of the 30 CRs initially adhering) CRs from 9 countries confirmed their participation and could timely reply to almost all questions. Only the Polish CRs group were not yet allowed to answer to the added 5 questions, as the National Polish Cancer Registry needed to be consulted about Ethical Committee approvals and privacy limit.
The Parma CR (Italy) could not answer to the five additional questions, due to internal re-organization procedures occurred after the WASABY Steering Committee. Also in this case, the information will be provided as soon as possible. Two more CRs joined WASABY after the GRELL Meeting in Trento (16th-18 May 2018): Alto Adige (Italy) and Navarra (Spain) and answered to the questionnaire. Also 17 French CRs joined the study after the GRELL Meeting, but they did not answered to the questionnaire, because they will be managed by a local responsible, who will organize the data collection and analysis (according to the WASABY procedures).
Afterwards, two more Polish CRs joined the study (Masovia and Podkarpackie) and, after the 2019 AIRTUM Meeting, also Trapani CR joined the study.

The CRs withdrawing their availability to join the study were: Hamburg and Munich (Germany), Firenze-Prato and Friuli Venezia-Giulia (Italy), Kracow and Lower Silesia (Poland). Afterwards, also the Belgium and Bremen (Germany) CRs left the study, due to internal organisation reasons.

Therefore, at date, a total of 44 CRs (plus the National Polish CR coordinating its local CRs) are involved in the study.

The current numbers of CRs and Nations officially adhering to WASABY fully satisfy the number limits expressed as objectives in the project. The WASABY Coordination Group contacted and stimulated the participation of more ENCR CRs profiting from the opportunity of national and international meetings involving CRs (e.g., GRELL Meeting on the 16th-18th of May 2018 in Trento, Italy).
Therefore, the list and number of participating CRs may still increase, even after the closing of the final list, expected for the end of August (WP4 Milestone M4.2 – Complete list of participating CRs).

Tab. 2a – Final list of CRs interested in participating to WASABY.

Nation	Cancer Registry	Geo-code level (main)	Geo-code linkage by	Reference population available at	Maps available at	SES
Germany	Schleswig-Holstein	Municipality	Cancer Registry	Cancer Registry by calendar year	Cancer Registry	Local deprivation index (until 2009)
Italy	Alto Adige	Census Block	Cancer Registry + WP4	Cancer Registry by calendar year	Cancer Registry/ISTAT	EDI 2001+national index
	Napoli 3 South	x, y coordinates	Cancer Registry + WP4	Cancer Registry by calendar year	Cancer Registry/ISTAT	EDI 2001+national and local index
	Palermo	x, y coordinates	Cancer Registry + WP4	Cancer Registry by calendar year	Cancer Registry/ISTAT	EDI 2001+national and local index
	Parma	Census Block	Cancer Registry	Cancer Registry by calendar year	ISTAT	EDI 2001+national index
	Ragusa	Zip Code/Census Block	Cancer Registry +WP4	Cancer Registry by calendar year	ISTAT	EDI 2001+national index
	Siracusa	Census Block	WP4	Cancer Registry by calendar year	ISTAT	EDI 2001+national index
	Trapani	Census Block	WP4	Cancer Registry by calendar year	ISTAT	EDI 2001+national index
	Trento	x, y coordinates	Cancer Registry + WP4	Cancer Registry by calendar year	Cancer Registry/ISTAT	EDI 2001+national and local index
	Umbria	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer Registry/ISTAT	EDI 2001+national and local index
	Varese	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer Registry/ISTAT	EDI 2001+national index
Lithuania	Lithuania	Eldership	Cancer Registry	Cancer Registry by calendar year	Lithuania Statistics Office	None
Poland	Greater poland	Municipality	Cancer Registry	Cancer Registry by calendar year	General Geographic Database of Poland	National and local deprivation index
	Kielce	Municipality	Cancer Registry	Cancer Registry by calendar year	General Geographic Database of Poland	National and local deprivation index
	Silesia	Municipality	Cancer Registry	Cancer Registry by calendar year	General Geographic Database of Poland	National and local deprivation index
	Masovia	Municipality	Cancer Registry	Cancer Registry by calendar year	General Geographic Database of Poland	National and local deprivation index
	Podkarpackie	Municipality	Cancer Registry	Cancer Registry by calendar year	General Geographic Database of Poland	National and local deprivation index
Portugal	Central Portugal	Parish	Cancer Registry	Cancer Registry by calendar year	INE	EDI 2001
	Northern Portugal	Parish	Cancer Registry	Cancer Registry by calendar year	INE	EDI 2001
Slovenia	Slovenia	x, y coordinates	Cancer Registry	Cancer Registry by calendar year	Cancer Registry	EDI (in development)
Spain	Basque Country	Municipality/Census Block	To be defined	Cancer Registry by sporadic calendar year	Cancer Registry	EDI 2001
	Castellon-Valencia	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer Registry	EDI 2001
	Girona	x, y coordinates	External paid resources	Cancer Registry by sporadic calendar year	Purchasable externally	EDI 2001
	Granada	x, y coordinates	External paid resources	Cancer Registry by calendar year	Cancer Registry	EDI 2001
	Murcia	Census Block	External paid resources	Cancer Registry by calendar year	Cancer Registry	EDI 2001
	Navarra	Census Block	To be defined	Cancer Registry by sporadic calendar year	To be defined	EDI 2001
UK	Northern Ireland	Zip Code	Cancer Registry	Cancer Registry by calendar year	Don't know	Local deprivation index

Legenda: ISTAT = Italian National Statistics Office; INE = National Institute of Statistics Portugal; EDI 2001 = European Deprivation Index 2001 (Guillaume et al., 2016).

Tab. 2b – Final list of French CRs interested in participating to WASABY.

Nation	Cancer Registry	Geo-code level (main)	Geo-code linkage by	Reference population available at	Maps available at	SES
France	Gironde	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Poitou-Charentes	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Loire-Atlantique et Vendée	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Haute Vienne	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Calvados	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Manche	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Somme	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Lille et sa région	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Bas-Rhin	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Haut-Rhin	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Doubs et Territoire de Belfort	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Cancers gynécologiques de Côte-d'Or	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Isère	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Hérault	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Tarn	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Guadeloupe	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index
	Martinique	Census Block	Cancer Registry	Cancer Registry by calendar year	Cancer RegistryFRANCIM	EDI 2001+national index

Legenda: FRANCIM = Réseau Français des Registres des Cancers; EDI 2001 = European Deprivation Index 2001 (Guillaume et al., 2016).

4. The results of the final Survey: characteristics of the confirmed participating Cancer Registries

The results of the preliminary survey showed different scenarios across the European CRs in terms of data availability and previous experiences in field of spatial analysis. Thus, WP leaders and Steering Committee decided to not to confine the study to one single standardized method for all CRs, but to apply different approaches for each country or CR, according to their available data.

The three main information collected through the questionnaire regarded the geographical level at which data are available, the existence of a socio-economic deprivation index usable by the CR and the presence of confidentiality issues requiring the involvement of a local Ethical Committee.

As for geographical level:

- Municipality: all 44 CRs.
- Zip Code (4 CRs): Ragusa, Kielce, Central Portugal, Northern Ireland. CRs providing completely geo-coded data (x & y) are also able to produce zip code information.
- Census tract or similar information (eldership, parish) (31 CRs): Alto Adige, Parma, Ragusa, Siracusa, Trapani, Umbria, Varese, Lithuania, Central Portugal, Northern Portugal, Basque Country, Castellon-Valencia, Murcia, Navarra and all the French CRs. CRs providing completely geo-coded data (x & y) also could be also able to produce Census tract information.
- X & Y coordinates (6 CRs): Napoli 3 South, Palermo, Trento, Slovenia, Girona, Granada.

As for the presence of a socio-economic deprivation index (Tab. 2):

- EDI was present in its 2001 version (updating), or in development, in 36 CRs, at Census tract level.
- National deprivation indices, at various geographical levels, were available for 15 CRs.
- Local deprivation indices, at various geographical level, were available for 10 CRs.
- Only the Lithuanian CR did not record a deprivation index.

As for differences in Ethical Committee approval:

- Fourteen CRs declared that INT Ethical Committee approval would be enough for their involvement in WASABY: Belgium, Napoli 3 South, Ragusa, Trapani, Varese, Lithuania, Central Portugal[2], Slovenia, Basque Country, Castellon-Valencia, Girona, Granada, Murcia, Northern Ireland.
- Six CRs declared their involvement is linked with the approval of local/specific Ethical Committees: Bremen, Schleswig-Holstein, Siracusa, Trento, Umbria, Northern Portugal,.
- A small number of CRs (5) were dealing with their referent national and/or local institutions, in order to devise if they need a specific Ethical Committee approval other than the INT one: Palermo, Parma, Greater Poland, Kielce, Silesia.

Another relevant information was the different time range for which the CRs can provide data for the study. In the following graphics (Graph. 1a & 1b), a comparison of the different period are shown. A synthesis of other relevant information was reported in Table 2.

Information by single CR were reported in the following paragraphs. Also the CRs which originally agreed to participate, filled the questionnaire and retired are reported.
Obviously, in the three-years lifespan of the project, some characteristics may change. In particular the geo-coding of the information could improve at a lower and more detailed geographical level (Census block or x & y coordinates) where not already available, thanks to specific technical intervention inside the WASABY framework.

[2] Central Portugal CR participation in the study was superiorly authorized by the Hospital administration.

Graph. 1 – Data availability from the CRs: time range.

CR	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Belgium																											
Bremen																											
Schleswig-Holstein																											
Alto Adige																											
Napoli 3 South																											
Palermo																											
Parma																											
Ragusa																											
Siracusa																											
Trapani																											
Trento																											
Umbria																											
Varese																											
Lithuania																											
Central Portugal																											
Northern Portugal																											
Slovenia																											
Basque Country																											
Castellon-Valencia																											
Girona																											
Granada																											
Murcia																											
Navarra																											
Northern Ireland																											
Greater Poland	**The 5 Polish CRs received the approval from the National Cancer Registry answer. They will process and analyse data by themselves, sharing only the results.**																										
Kielce																											
Silesia																											
Masovia																											
Podkarpackie																											

CR	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Gironde																											
Poitou-Charentes																											
Loire-Atlantique et Vendée																											
Haute Vienne																											
Calvados																											
Manche																											
Somme																											
Lille et sa région																											
Bas-Rhin																											
Haut-Rhin																											
Doubs et Territoire de Belfort																											
Cancers gynécologiques de Côte-d'Or																											
Isère																											
Hérault																											
Tarn																											
Guadeloupe																											
Martinique																											

4.1 Belgium Cancer Registry

CR's Director: Liesbet Van Eycken (elizabeth.vaneycken@kankerregister.org). **WASABY referents**: Julie Francart (Julie.Francart@registreducancer.org) and Kris Henau (Kris.Henau@kankerregister.org).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality.
SU dimension (average number of people)	18931 (range: 58 – 510610).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI in development.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2004-2016.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes (Municipality).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Concerning the privacy and confidentiality aspects, the requirements in Belgium are very strict. It is possible that a specific request must be submitted to the Privacy Commission to obtain authorization before transferring the data. These discussions will take some time and will delay the final response.
Detail limits in publishing maps (see confidentiality above).	No limits in mapping municipality data, if Privacy Commission will allow to send data.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.2 Bremen Cancer Registry

CR's Director and WASABY referent: Sabine Luttman (luttmann@bips.uni-bremen.de).

Question	Provided information
Geo-code level (smallest unit - SU).	Urban District (Bremen).
SU dimension (average number of people)	66189.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	Local index (until 2009) and EDI in development.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2004-2013.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	For aggregated data on municipality unit only an approval from local government, which usually takes 3-4 weeks. On small area unit (Urban District) it would need to aggregate on a 10 year time period and an approval from local Ethic Committee is needed. Very small number (< 5) in some small areas are expected (leading to potential data anonymization).
Detail limits in publishing maps (see confidentiality above).	As above.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	As above

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.3 Schleswig-Holstein Cancer Registry

CR's Director: Alexander Katalinic (Alexander.Katalinic@uksh.de). **WASABY referent**: Ron Pritzkuleit (Ron.Pritzkuleit@uksh.de).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality.
SU dimension (average number of people)	2474 (range: 5 – 78263).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	Local index (until 2009) and EDI in development.
Previous studies in the area, connecting cancer outcomes and environmental issues.	No experiences.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2000-2015.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	Permanent small changes but population data and incidence data can be converted to a unique area status/mapping status.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Assuming that the data used is either anonymized or aggregated, in both cases there should be no problems with confidentiality. This should be pointed out for EC approval.
Detail limits in publishing maps (see confidentiality above).	There is no problem to publish rates or smoothed rates. It could be a problem for small municipalities for numbers.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.4 Alto Adige Cancer Registry

CR's Director: Guido Mazzoleni. **WASABY referent**: Andreas Bultako (andreasklaus.bulatko@sabes.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality.
SU dimension (average number of people)	Globally 526000 people .
Geo-code linkage by:	To be defined.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI and national deprivation indices.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study (including breast cancer in young women).
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1995-2013.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Problems might arise.
Detail limits in publishing maps (see confidentiality above).	Municipality level.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Local Ethical Committee approval must be informed.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.5 Napoli 3 South Cancer Registry

CR's Director: Mario Fusco (mariofusco2@virgilio.it). **WASABY referent**: Francesca Maria Vitale (mafravi86@libero.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block.
SU dimension (average number of people)	170 (range: 1 – 3386).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI and national and local deprivation indices.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1997-2015.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2001, 2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Problems will not arise.
Detail limits in publishing maps (see confidentiality above).	Problems will not arise.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.6 Palermo Cancer Registry

CR's Director: Francesco Vitale (francesco.vitale@unipa.it). **WASABY referents:** Rosanna Cusimano (rosanna.cusimano@unipa.it), Walter Mazzucco (walter.mazzucco@unipa.it) and Maurizio Zarcone (registrotumoripalermo@unipa.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (probably Census block).
SU dimension (average number of people)	Census block: 170 (range: 1 – 3386).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI and national and local deprivation indices.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2001-2015.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2001, 2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Still to be defined.
Detail limits in publishing maps (see confidentiality above).	Still to be defined.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Still to be defined.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.7 Parma Cancer Registry

CR's Director: Maria Michiara (michiara@ao.pr.it). **WASABY referent:** Paolo Sgargi (psgargi@ao.pr.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block (but only for Parma municipality).
SU dimension (average number of people)	170 (range: 1 – 3386).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	ISTAT.
SES and deprivation index.	EDI and national deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2001-2013.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2001, 2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Still to be defined.
Detail limits in publishing maps (see confidentiality above).	Still to be defined.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Still to be defined.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.8 Ragusa Cancer Registry

CR's Director and WASABY referent: Rosario Tumino (rtumino@tin.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Zip Code and x & y coordinates.
SU dimension (average number of people)	Zip Code: 26766.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	ISTAT.
SES and deprivation index.	EDI and national deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1998-2015.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes (x &y coordinates).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Problems will not arise.
Detail limits in publishing maps (see confidentiality above).	Problems will not arise.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.9 Siracusa Cancer Registry

CR's Director: Anselmo Madeddu (rtp@ausl8.siracusa.it). **WASABY referent:** Francesco Tisano (rtp@asp.sr.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (but they will move towards Census block).
SU dimension (average number of people)	Census block: 170 (range: 1 – 3386).
Geo-code linkage by:	They must identify someone inside or outside the CR to perform geo-coding.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	ISTAT.
SES and deprivation index.	EDI and national deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2004-2013.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Local Ethical Committee approval is needed.
Detail limits in publishing maps (see confidentiality above).	Problems will not arise.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.10 Trapani Cancer Registry

CR's Director: Giuseppa Candela (candelag@inwind.it). **WASABY referent:** Tiziana Scuderi (trusyit@yahoo.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (but they will move towards Census block).
SU dimension (average number of people)	Census block: 170 (range: 1 – 3386).
Geo-code linkage by:	They must identify someone inside or outside the CR to perform geo-coding.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	ISTAT.
SES and deprivation index.	EDI and national deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2002-2011.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Problems will not arise.
Detail limits in publishing maps (see confidentiality above).	Problems will not arise.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.11 Trento Cancer Registry

CR's Director and WASABY referent: Silvano Piffer (Silvano.Piffer@apss.tn.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block.
SU dimension (average number of people)	170 (range: 1 – 3386).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI and national and local deprivation indices.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2003-2012.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Local Ethical Committee approval is needed.
Detail limits in publishing maps (see confidentiality above).	There is a problem of confidentiality, given the municipal disaggregation of Trentino and the consequent possibility of identifying cases for municipalities with few inhabitants.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Local Ethical Committee approval is needed.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.12 Umbria Cancer Registry

CR's Director and WASABY referent: Fabrizio Stracci (fabrizio.stracci@unipg.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block and x & y coordinates.
SU dimension (average number of people)	Census block: 170 (range: 1 – 3386).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI and national and local deprivation indices.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2001-2015.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2001 and 2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Confidentiality issues in conferring nominative and coordinated data without approval of the Ethical Committee and of the Region.
Detail limits in publishing maps (see confidentiality above).	No limits.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Local Ethical Committee approval.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.13 Varese Cancer Registry

CR's Director: Giovanna Tagliabue (Giovanna.Tagliabue@istitutotumori.mi.it). **WASABY referents:** Paolo Contiero (Paolo.Contiero@istitutotumori.mi.it) and Giovanna Tagliabue (Giovanna.Tagliabue@istitutotumori.mi.it).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block and x & y coordinates.
SU dimension (average number of people)	170 (range: 1 – 3386).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI and national deprivation indices.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, correlation with cancer was studied, also breast cancer was considered, but not yet specifically in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1990-2012.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2001 and 2011 (Census data collection).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Problems will not arise.
Detail limits in publishing maps (see confidentiality above).	Problems will not arise.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.14 Lithuania Cancer Registry

CR's Director and WASABY referent: Ieva Vincerževskien (ieva.vincerzevskiene@nvi.lt).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality.
SU dimension (average number of people)	58117 (range: 2400 – 554300).
Geo-code linkage by:	They must identify someone inside or outside the CR to perform geo-coding.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Lithuania National Statistics Office.
SES and deprivation index.	No deprivation index is available.
Previous studies in the area, connecting cancer outcomes and environmental issues.	None.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1990-2012.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes (municipality).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Problems will not arise.
Detail limits in publishing maps (see confidentiality above).	Problems will not arise.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.15 Greater Poland Cancer Registry

CR's Director: Maciej Trojanowski (maciej.trojanowski@wco.pl). **WASABY referent:** Łukasz Taraszkiewicz (lukasz.taraszkiewicz@wco.pl).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (LAU-2).
SU dimension (average number of people)	15,405 (as of 2016).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Administrative maps can be extracted from the General Geographic Database of Poland led by the Poland's Head Office of Geodesy and Cartography.
SES and deprivation index.	National and/or local deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with breast cancer has not been studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2007–2016.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	Small changes of some municipality borders (<20 municipalities).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems are expected.
Detail limits in publishing maps (see confidentiality above).	To avoid identification of single cancer patients living in areas with low population density, one proposes to map interval incidence data (with the first interval e.g. 0–3) or to aggregate data from the entire analysis period for map presentation.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough. Local Ethical Committee has been also informed about the project.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.16 Kielce Cancer Registry

CR's Director: Stanislaw Gozdz (stanislaw.gozdz@onkol.kielce.pl). **WASABY referent:** Ryszard Mezyk (Ryszard.Mezyk@onkol.kielce.pl).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (LAU-2).
SU dimension (average number of people)	12,283 (as of 2016).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Administrative maps can be extracted from the General Geographic Database of Poland led by the Poland's Head Office of Geodesy and Cartography.
SES and deprivation index.	National and/or local deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with breast cancer has not been studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2007–2016.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	Small changes of some municipality borders (<5 municipalities).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems are expected.
Detail limits in publishing maps (see confidentiality above).	To avoid identification of single cancer patients living in areas with low population density, one proposes to map interval incidence data (with the first interval e.g. 0–3) or to aggregate data from the entire analysis period for map presentation.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough. Local Ethical Committee has been also informed about the project.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.17 Silesia Cancer Registry

CR's Director and WASABY referent: Marcin Motnyk (Marcin.Motnyk@io.gliwice.pl).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (LAU-2).
SU dimension (average number of people)	27,300 (as of 2016).
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Administrative maps can be extracted from the General Geographic Database of Poland led by the Poland's Head Office of Geodesy and Cartography.
SES and deprivation index.	National and/or local deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with breast cancer has not been studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2007–2016.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	Small changes of some municipality borders (<10 municipalities).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems are expected.
Detail limits in publishing maps (see confidentiality above).	To avoid identification of single cancer patients living in areas with low population density, one proposes to map interval incidence data (with the first interval e.g. 0–3) or to aggregate data from the entire analysis period for map presentation.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough. Local Ethical Committee has been also informed about the project.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.18 Central Portugal Cancer Registry

CR's Director: Margarida Ornelas. **WASABY referent:** Joana Antunes Lima Bastos (3603@ipocoimbra.min-saude.pt).

Question	Provided information
Geo-code level (smallest unit - SU).	Zip code for this specific study.
SU dimension (average number of people)	To be defined.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	None.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1998-2011.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	The last census was in 2011 and the other one before was in 2001. There was a reorganization of Portuguese smallest sub-area of residence (freguesia) in 2012.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	No confidentiality problems.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.19 Northern Portugal Cancer Registry

CR's Director: Maria José Bento (mjbento@ipoporto.min-saude.pt). **WASABY referent:** Luis Antunes (luis.antunes@ipoporto.min-saude.pt).

Question	Provided information
Geo-code level (smallest unit - SU).	Parish.
SU dimension (average number of people)	1088.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	INE.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2003-2012.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	Data will be aggregated by parish (smallest geographical area available). There have been some changes in the number and composition of these units over time but distribution of parishes used in the census 2011 will be used.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	Local Ethical Committee approval.
Detail limits in publishing maps (see confidentiality above).	Local Ethical Committee approval.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Local Ethical Committee approval.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.20 Slovenia Cancer Registry

CR's Director: Maja Primic Žakelj (MZakelj@onko-i.si). **WASABY referents:** Tina Žagar (TZagar@onko-i.si) and Vesna Zadnik (vzadnik@onko-i.si).

Question	Provided information
Geo-code level (smallest unit - SU).	X & Y coordinates.
SU dimension (average number of people)	It depends on the area radius around the coordinates.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1998-2015.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes (x & y coordinates).
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	No confidentiality problems.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.21 Basque Country Cancer Registry

CR's Director: Arantza Lopez De Munain Marques (arantza-lopez@euskadi.eus). **WASABY referent:** Nerea Larrañaga (n-larranaga@euskadi.eus).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality. Census block available from 2014.
SU dimension (average number of people)	Census block: 1500.
Geo-code linkage by:	They must identify someone inside or outside the CR to perform geo-coding.
Reference population available at:	Cancer Registry by sporadic calendar year.
Maps available at:	To be identified by CR.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Correlation with cancer is/it has been under study, yet never with breast cancer in young women.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2005-2015 (Census bock: 2014-2015).
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes in the study period.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	No confidentiality problems.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.22 Castellon-Valencia Cancer Registry

CR's Director: Consol Sabater Gregori (sabater_congre@gva.es). **WASABY referents:** Paloma Botella (botella_pal@gva.es) and Fernando Almela (almela_fer@gva.es).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block.
SU dimension (average number of people)	1500.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2005-2014.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2011 Census.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	No confidentiality problems.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.23 Girona Cancer Registry

CR's Director: Rafael Marcos-Gragera (rmarcos@iconcologia.net). **WASABY referent:** Marc Saez (marc.saez@udg.edu).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block and x & y coordinates.
SU dimension (average number of people)	1500.
Geo-code linkage by:	External paid resources.
Reference population available at:	Cancer Registry by sporadic calendar year.
Maps available at:	Purchasable externally.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1994-2014.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	2011 Census.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	No confidentiality problems.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.24 Granada Cancer Registry

CR's Director and WASABY referent: María José Sánchez Pérez (mariajose.sanchez.easp@juntadeandalucia.es).

Question	Provided information
Geo-code level (smallest unit - SU).	Municipality (x & y coordinates could be attempted).
SU dimension (average number of people)	5473 (range: 134 – 237540).
Geo-code linkage by:	External paid resources.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2004-2013.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes in the study period.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	No confidentiality problems.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.25 Murcia Cancer Registry

CR's Director: María Dolores Chirlaque (mdolores.chirlaque@carm.es).. **WASABY referent:** María Dolores Chirlaque López (mdolores.chirlaque@carm.es) and Mónica Ballesta (monica.ballesta@carm.es).

Question	Provided information
Geo-code level (smallest unit - SU).	Census block.
SU dimension (average number of people)	1500.
Geo-code linkage by:	Cancer Registry and External paid resources.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	Cancer Registry.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Environmental studies were performed, but correlation with cancer was never studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1996-2012.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes in the study period.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	In small areas only is possible to publish more than 5 cases.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.26 Navarra Cancer Registry

CR's Director: Eva Ardanaz (me.ardanaz.aicua@cfnavarra.es). **WASABY referent:** Marcela Guevara (mp.guevara.eslava@navarra.es).

Question	Provided information
Geo-code level (smallest unit - SU).	Basic Health zones and zx & y coordinates.
SU dimension (average number of people)	11000 (varying from 1300 to 22000 people).
Geo-code linkage by:	To be defined.
Reference population available at:	Cancer Registry by sporadic calendar year.
Maps available at:	To be defined.
SES and deprivation index.	EDI.
Previous studies in the area, connecting cancer outcomes and environmental issues.	Yes, but correlation with cancer was never studied.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	2009-2012.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	To be defined.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	There could be confidentiality problems for very small geographic areas, and also for rare cases.
Detail limits in publishing maps (see confidentiality above).	There could be confidentiality problems for very small geographic areas, and also for rare cases.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	Local Ethical Committee approval.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.27 Northern Ireland Cancer Registry

CR's Director and WASABY referent: Anna Gavin (A.Gavin@qub.ac.uk).

Question	Provided information
Geo-code level (smallest unit - SU).	Zip code.
SU dimension (average number of people)	4342.
Geo-code linkage by:	Cancer Registry.
Reference population available at:	Cancer Registry by calendar year.
Maps available at:	External provider to be identified
SES and deprivation index.	Local deprivation index.
Previous studies in the area, connecting cancer outcomes and environmental issues.	None.
The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).	1993-2016.
The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).	No changes in the study period.
Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.	No confidentiality problems.
Detail limits in publishing maps (see confidentiality above).	There may be some copyrights to be acknowledged when publishing maps.
Do you think your institute will require a specific EC approval for the participation in this project or will the INT EC approval be sufficient?	INT Ethical Committee approval is enough.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

5. Conclusions and next steps

The survey enabled to collect information on data available each CR interested to participate. Based on the findings, two major deliverables will be produced: a protocol for CR on data collection on breast cancer and a model to map cancer incidence.
CRs will need to provide the list of first invasive breast cancer cases (coded as C50 according to the ICD-10) diagnosed during a specific period (to be defined separately for each participating CR, e.g. since 1990), together with age at diagnosis (or 5-year age groups), morphology and data on the place of residence at the time of diagnosis (exact x and y coordinates or smallest possible sub-area of residence).
More specifically:

- Residence at smallest possible sub-area level for each cancer case, on the date of diagnosis. Some CRs already have this detail and no additional work is needed. CRs that now cannot provide such information will be helped with operative intervention (in terms of education and practical actions) by WP4, after an exploration aimed to evaluate the efficacy of the intervention.
- Female population data by 5-year age groups, calendar year within time period and sub-area on which the incidence data would be estimated. Sub-areas refer to the smallest geographical area for which required data are available and may differ between countries.
- Shape files for the area covered by CR and sub-areas are usually provided by the national mapping authority. Centroids of the sub-areas are calculated from GIS software from given shape files.
- Pollutants indicators by sub-area (for those CRs involved in the pilot environmental study, as reported in the WASABY Project).

A detailed description of the actions that every CR is expected to accomplish the WASABY tasks, will be presented in the Protocol for CR data collection and Ethical Committees (Deliverable D4.2) due at the end of May 2018 (M5).

Moreover, a complete list of participating CRs will be presented at the end of August (Milestone M4.2), possibly including additional CRs, contacted between May and July 2018.

Indeed, a key element in the reflection of EDI construction is the smallest geographical entity for which census data are available in the country concerned. This information will give for each country the scale to which the incidence data must be geo-coded in order to account for deprivation. WP5 will try to retrieve this information from the relevant institutions in each country. Only once this information is available, the management of the geo-coding work in each country will be possible. To have optimal conditions for the EDI construction in each country, information on all the geographical area available for each country (administrative and non-administrative) and the link between them is also needed. So, WP5 proposes to elaborate a technical point on census level availability for each country and then send a questionnaire to WASABY contacts involved in EDI construction to have information about geographical unit in their country. That will give for each country the scale of EDI construction which will be the largest scale needed for geo-coding cases to account for deprivation in further analyses.

References

Guillaume E, Pornet C, Dejardin O, Launay L, Lillini R, Vercelli M, Marí-Dell'Olmo M, Fernández Fontelo A, Borrell C, Ribeiro AI, Pina MF, Mayer A, Delpierre C, Rachet B, Launoy G. (2016) *Development of a cross-cultural deprivation index in five European countries.* Journal of Epidemiology and Community Health; 70(5):493-499.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

PROTOCOL FOR
MAPPING OF BREAST CANCER RISK FOR THE WASABY PROJECT
V1 – 23rd February 2018

1. INTRODUCTION

The WASABY project (WAter & Soil contamination and Awareness on Breast cancer risk in Young women) focuses on the geographical analysis of population based cancer incidence data in connection with environmental factors, using breast cancer and water/soil contamination as an exemplification replicable to other cancer sites. The following table presents the participants to the project:

Participants	Acronym	Country	Work Packages (WP)
FONDAZIONE IRCCS ISTITUTO NAZIONALE DEI TUMORI	INT	Italy	WP1: Coordination of the project WP4: Data management WP7: Environmental risk factors & breast cancer
ASSOCIATION EUROPEENNE DES LIGUES CONTRE LE CANCER ASBL	ECL	Belgium	WP2: Dissemination of the project
UNIVERSITÄT ZU LÜBECK	GER	Germany	WP3: Evaluation of the project
UNIVERSITE DE CAEN NORMANDIE	FRA	France	WP5: Deprivation indexes
ONKOLOSKI INSTITUT LJUBLJANA	SLO	Slovenia	WP6: Methods and analysis

The present protocol focuses only on the WASABY project activity of spatial analysis for breast cancer risk in the participating cancer registries. The list of potential participating cancer registries (CRs) is in Annex 1. The project requires to have at least 15 CRs from 6 European countries. Formally a CR will be considered a participating CR when a) it demonstrates to be able to geo-code own cases and b) an ethical committee will allow it to participate in the project.

The present protocol is prepared as *vademecum* for each CR interested to participate in the project. WASABY allows CRs to define incidence years and type of geographic data, for this reason different methods may be applied according to original data received.

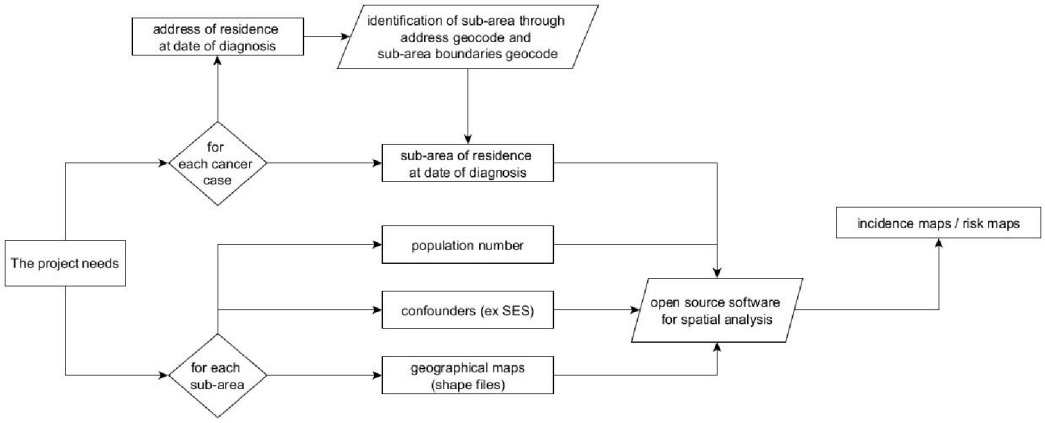
Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

SINTHESIS OF THE PROJECT

Since individuals represent the basic unit of spatial analysis in cancer research, each CR shall assign geographic information (exact x and y coordinates or smallest possible sub-area of residence (SU)) to every breast cancer case corresponding to the location of their place of residence.

The following schema shows the process of data collection for the present protocol.



According to the schema above, the protocol steps are:

1. Each participating CR is required to provide information on breast cancer cases (coded as C50 according to the ICD-10) diagnosed during a specific ten-year period (to be defined separately for each participating CR, e.g. 2001-2010), together with age at diagnosis (or 5-year age groups), morphology and data on the place of residence at the time of diagnosis (exact x and y coordinates or SU). See Annex 2 for details on data to be collected
2. Socio Economic Status (SES) data will be collected as main confounder in the spatial analysis: National or European Deprivation indexes by SU will be utilized. See Annex 3 for details on SES and other potential confounders
3. Maps of incidence will be estimated in order to identify CR SU characterized by higher-than-CR average rates. See Annex 4 for methods to be applied according to data available in each CR.

Before the start of the data collection, every CR is required provide the following general information, which will be added to the ones collected with the preliminary survey:

- The calendar years by which the CR can provide incidence data at the most disaggregated geographic level (exact x and y coordinates or SU).
- The calendar year by which such geographic level has changed (e.g., census tract changes between two different Census data collection).
- Any confidentiality problems likely to arise if/when pursuing the approval to the Ethical Committees, locally.
- Detail limits in publishing maps (see confidentiality above).

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Data storage

Two modalities of data storage can be envisaged:

- OPTION 1
  - Data will be centrally stored at FONDAZIONE IRCCS ISTITUTO NAZIONALE DEI TUMORI and, only for the selected number of cancer registries involved in WP6 analyses, data will be shared with the ONKOLOSKI INSTITUT LJUBLJANA
  - Data will be stored individually (but anonymously). If a CR is to send breast cancer cases by SU, data will be stored at aggregated level
  - Data will be stored in a dedicated server not connected to the web, and according to the standard requirements for data security
  - Data handling will be conform with the EC General Data Protection Regulation (2016/679)
- OPTION 2
  - Only results of the analysis (performed by the CR) will be shared in the WASABY project. This will be the case of CRs with the entire population geocoded.

4. INT contacts for data collection

Every information request and submission, in particular regarding the four points above, which update the questionnaire data, must be addressed to:

roberto.lillini@istitutotumori.mi.it (Roberto Lillini)

and CC:

lifetable@istitutotumori.mi.it (Paolo Baili)

5. Publication Policy of the entire project WASABY

All publications performed in the WASABY context must mention the WASABY Working Group. A suitable authorship formula being: Authors A, B, C, ... and the WASABY Working Group, with all members listed in a footnote or appendix to the article.

The WASABY Working Group will be realized with the following members:

- All members of the Steering Committee (SC)
- All members of the Management Support Team (MST)
- Up to two members of each Partner indicated in the introduction (in addition to those included in the SC and MST)
- Up to two members of each Cancer Registry participating in WASABY
- All experts actively participating in the work packages of the project
- Up to one member for each participating area working in geocoding activities (unless included in the previous points)

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 1 – TENTATIVE LIST OF PARTICIPATING CANCER REGISTRIES

Nation	Cancer Registry
Belgium	Belgium
Germany	Bremen
Germany	Schleswig-Holstein
Italy	Napoli 3 South
Italy	Palermo
Italy	Parma
Italy	Ragusa
Italy	Siracusa
Italy	Trento
Italy	Umbria
Italy	Varese
Lithuania	Lithuania
Poland	Greater Poland
Poland	Kracow
Poland	Kielce
Poland	Silesia
Portugal	Central Portugal
Portugal	Northern Portugal
Slovenia	Slovenia
Spain	Basque Country
Spain	Castellon-Valencia
Spain	Girona
Spain	Granada
Spain	Murcia
UK	Northern Ireland

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 2 – DATA TO BE COLLECTED FOR EACH CANCER REGISTRY

FILE WITH BREAST CANCER CASES AND GEOGRAPHIC DATA

Primary invasive female breast cancer (ICD9 174*, ICD10 C50*), selected from cancer registries data during a specific ten years period (ex: 2001 to 2010) are included in the project. It is mandatory to collect data with age at diagnosis less than 50 years of age, while it is not mandatory to collect data for all ages. Synchronous and metachronous breast cancer cases must be counted once. Cancer registration criteria must follow European Network of Cancer Registries (ENCR) rules. Residence addresses at diagnosis retrieved from the National or local Security system or from the personal data reference of each registry will be collected.

Data can be collected in two different modalities.

OPTION 1 – individual level

		Variable name	Description	Data type	Mandatory
BREAST CANCER VARIABLES		CR	Cancer Registry name	Alphanumeric variable	Yes
		PATIENT_ID	Patient identification code assigned by Cancer Registry.	Numeric / Alphanumeric variable	Yes
		DATE OF DIAGNOSIS	Incidence date based on histological or cytological confirmation of the malignancy	DD/MM/YYYY	Yes
		DATE OF BIRTH	Date of birth of the patient	DD/MM/YYYY	Yes (one of the two variables)
		AGE	Age at diagnosis	Numeric variable	
		ICDO3_M	ICDO3 morphology code of incident case	Alphanumeric variable	Yes
		SUBTYPE_ER	Estrogen Receptor value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_PGR	Progesterone Receptor value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_HER2	HER-2 value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_KI67	KI-67 value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_FISH	FISH value at diagnosis	Numeric / Alphanumeric variable	No
GEOGRAPHIC VARIABLES	OPTION A	X	Longitude coordinate referred to the address where the patient was residing at the moment of the breast cancer diagnosis	Numeric variable	Yes, data from one option
		Y	Latitude coordinate referred to the address where the patient was residing at the moment of the breast cancer diagnosis	Numeric variable	
		Reference	The coordinate system used for X and Y: UTM WGS84 32N vs. UTM ED 1950 32N	Alphanumeric variable	
	OPTION B	SU	Smallest administrative unit (SU) where the patient was residing at the moment of the breast cancer diagnosis	Alphanumeric variable	
	OPTION C	MUNICIPALITY_ CODE	Code of the Municipality where the patient was residing at the moment of the breast cancer diagnosis	Alphanumeric variable	
		MUNICIPALITY	Name of the Municipality where the patient was residing at the moment of the breast cancer diagnosis	Alphanumeric variable	

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

OPTION 2 – aggregated level

		Variable name	Description	Data type	Mandatory
BREAST CANCER VARIABLES		CR	Cancer Registry name	Alphanumeric variable	Yes
		YEAR DIAGNOSIS	Incidence year based on histological or cytological confirmation of the malignancy	Numeric variable	Yes
		AGE	Age class at diagnosis	Alphanumeric variable	Yes
		ICDO3_M	ICDO3 morphology code of incident case	Alphanumeric variable	Yes
		SUBTYPE_ER	Estrogen Receptor value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_PGR	Progesterone Receptor value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_HER2	HER-2 value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_KI67	KI-67 value at diagnosis	Numeric / Alphanumeric variable	No
		SUBTYPE_FISH	FISH value at diagnosis	Numeric / Alphanumeric variable	No
GEOGRAPHIC VARIABLES	OPTION B	SU	Smallest administrative unit (SU) where the patient was residing at the moment of the breast cancer diagnosis	Alphanumeric variable	Yes, data from one option
	OPTION C	MUNICIPALITY_ CODE	Code of the Municipality where the patient was residing at the moment of the breast cancer diagnosis	Alphanumeric variable	
		MUNICIPALITY	Name of the Municipality where the patient was residing at the moment of the breast cancer diagnosis	Alphanumeric variable	
DATA		NR_CASES	Number of primary invasive female breast cancer by all the previous variables	Numeric variable	Yes

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

POPULATION FILES

For every CR, WASABY needs the reference population at the same geographic level on which that CR intends to study the incident cases. More specifically, the population files must contain the female population data by 5-year age groups, calendar year within time period and SU (sub-areas refer to the smallest geographical area for which required data are available and may be different across countries).

All the variables are mandatory.

Variable name	Description	Data type
CR	Cancer Registry name	Alphanumeric variable
AGE_CLASS	5-year age class	Numeric/Alphanumeric variable
YEAR	Calendar year	Numeric/Alphanumeric variable
REF_DATE	Reference date of population data (1st Jan, 31st Dec, ecc)	Date/Alphanumeric variable
SU	MUNICIPALITY_CODE or SU indicated in the file with geographic data (see pages 5 or 6)	Alphanumeric variable
POP	Female population by 5-year age groups, calendar year within time period and sub-area on which the incidence data would be estimated	Numeric

SHAPEFILES

For every CR, WASABY needs a complete shapefile of the geographic area covered by its activity. The shapefile format is a digital vector storage format for storing geometric location and associated attribute information. It consists of a collection of files with a common filename prefix (e.g., Varese.shp, Varese.dbf, Varese.shx), stored in the same directory, with mandatory and optional files.

Mandatory files:

File name	Description	Data type
(CR area).shp	Shape format; the feature geometry itself	Alphanumeric
(CR area).shx	Shape index format; a positional index of the feature geometry to allow seeking forwards and backwards quickly	Alphanumeric
(CR area).dbf	Attribute format; columnar attributes for each shape, in dBase IV format	Alphanumeric

Files must be combined with information on calendar years of validity (in case of administrative changes of SU in the incidence years studied).

Other optional files, regarding spatial features not reported in the .dbf file, can be added but are not needed for a correct representation.

In the .dbf file an information about the minimum geo-coding level must be reported (i.e., census block, municipality, etc.)

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 3 – CONFOUNDERS

Socio economic status (SES) and other confounders: Deprivation index

Since this study includes different European countries, it is important that measurement of socioeconomic deprivation be comparable or at least transferable between different European countries, despite their socio-cultural differences, to improve the comparability and reproducibility across countries. The European Deprivation Index (EDI) measures the social environment in a comparable manner across countries, despite the differences in the census variables available, and to incorporate the social and cultural specificities of each country concerned. The ecological deprivation indices are built according to shared methodological principles, by selecting fundamental needs associated with both objective and subjective poverty, and they use the same theoretical concept of relative deprivation using a European survey dedicated to relative deprivation (Eu-Silc) regularly conducted on national samples from the all European countries. The method for constructing national versions of this EDI is described in different papers [Pornet C, JECH 2012; Guillaume E, JECH, 2016] and national versions of EDI are already available for 5 European countries (Italy, Portugal, Spain, England and France). This index is based on two elements:

- The European survey on deprivation European Union Statistics on Income and Living Conditions (EU-SILC) is a cross-sectional and longitudinal sample survey providing data on income, poverty, social exclusion and living conditions in the European Union. From these data, the statistical office of the European Union (Eurostat—http://ec.europa.eu/eurostat/web/main) produces a European standardized questionnaire that is specifically designed to study deprivation. It consists of nine questions, common to European Union members, evaluating needs that directly or indirectly induce financial inability. For each European Union member, the sum of weights for the sample design and the response rate to a national questionnaire were tailored on the basis of the national population size. All analyses were weighted for non-response and adjusted for sample design, to ensure the representativeness of the results for each member.
- The ecological data of the national population censuses. Ecological data came from the last exhaustive national population censuses, which were conducted in 2001 for Italy (Italian National Institute of Statistics: ISTAT), Portugal (National Institute of Statistics: INE), Spain (National Institute of Statistics: INE) and England (Office for National Statistics: ONS), and, in 1999, for France (National Institute for Statistics and Economic Studies: INSEE). To minimize the unavoidable ecological bias as much as possible, the smallest area for which census data were available was identified.

Also in this case the efforts performed are for the number of participating countries and not for the number of participating CRs. As already mentioned, at the time of writing this methodology has been developed for 5 countries but it may replicable in other European Union member states. Therefore, we will evaluate the construction of the EDI for countries with areas covered by participating cancer registries. In countries without available data to construct the EDI the collection of national deprivation indexes will be envisaged.

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Other confounders

Individual factors, e.g. ethnicity, family history, age, reproductive factors, alcohol intake, weight, physical activity, hormone therapy and oral contraceptives, have been found to influence the risk of breast cancer. Adherence to organized screening programmes in areas covered by cancer registries, lead to an increment of incidence in those areas [Pacelli, Eur J Public Health, 2014]; such information, however, is not available at individual level. Where possible, information on adherence to organized cancer screening is to be collected at SU level. If data are collected only for ages <50 screening adherence is not required.

The file with confounders is structured this way:

Variable name	Description	Data type	Mandatory
COUNTRY	Country name	Alphanumeric variable	Yes
SUB_AREA	MUNICIPALITY_CODE or SU indicated in the file with geographic data (see pages 5 or 6)	Alphanumeric variable	Yes
SES_SCALE	European Deprivation Index or specific national deprivation indices (according to the availability in the specific CR) by SU of incidence data. This is a scale variable	Numeric - Scale	Yes
SES_ORDINAL	European Deprivation Index or specific national deprivation indices (according to the availability in the specific CR) by SU of incidence data, classified by deprivation groups. This is an ordinal variable	Numeric - Ordinal	Yes
SCR_ADH	% of screening adherence by SU of incidence data.	Numeric - Scale	No

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 4 – METHODS

Identification of risk areas across European Countries Background

We will conduct this project using open source GIS software such as QGis [http://www.qgis.org/en/site/]. Furthermore, a single ArcGis desktop 10.0 license is also available for INT group and it will be used during the project to improve, if necessary, geographical and spatial analysis. GIS system software allows users to create maps with many layers (raster or vector) using different map projections. The vector data is stored as either point, line, or polygon-feature. Different kinds of raster images are supported, and the software can georeference images. Maps can be assembled in different formats and for different uses.

Spatial analysis

When large spatial units are used, the heterogeneity of exposure and different population characteristics can be missed. On the other hand, in small spatial units, the number of cancer cases is usually low and analysing the observed spatial pattern proves to be inefficient, as the population base, from which these cases arise, is often very low too. This can lead to unstable and misleading estimates of the true value. Modern approaches to relative risk estimation often rely on smoothing methods. The basic idea of mapping smoothed estimates is to borrow information from neighbouring regions to produce more stable and less noisy estimate associated with each geographical area and thus separate out the spatial pattern from the noise [Waller LA, Applied Spatial Statistics for Public Health Data, Wiley, NJ, 2004]. Taking into account these considerations we perform different statistical methods according to the available data, as follows:

- Estimate of a census block level breast cancer incidence risk using Generalised Additive Models (GAMs), a form of non-parametric or semi-parametric regression offering the possibility to analyse contextual data while adjusting for covariates and taking into account spatial autocorrelation [Woods SN, Chapman and Hall, USA 2006]. This model takes into account the spatial dependence of the data and the incidence rate variability that is due to the small number of events per geographic unit, by using a locally weighted regression smoother to account for geographic location as a possible predictor of incidence rate [Webster T, Env Health Persp, 2008]. With this model it is possible to estimate the relative risk by adjusting for covariates.

- Estimate of a census block relative breast cancer risk, by using the Besag, York and Mollié (BYM) model [Besag J, Ann Inst Stat Math, 1991], since it assumes the existence of two sources of extra variation, one spatial and the other non-spatial. The BYM model can be specified as a generalised linear mixed model (GLMM) with Poisson response variables, and considering the expected cases as an offset. The non-spatial random effect, also called heterogeneity, is usually assumed to be distributed with zero mean and constant variance. For the random effect, which captures spatial variability, a conditional autoregressive (CAR) [Clayton DG, Int J Epidem, 1993] model is used. The BYM model enables us to obtain smoothed estimates in each sub-area and, on the other hand, to estimate the effects of possible explanatory variables, such as the deprivation index.

Open source software is used for data manipulation and statistical analyses such as R and WinBUGS.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 10 – M4.1 – ETHICAL COMMITTEES APPROVAL

INT and local Ethical Committee approvals by Cancer Registry

Background

Some of the Cancer Registries (CRs) collaborators of the project declared that an approvals by local Ethical Committees (ECs) was due, in order to be allowed to participate to the study, while for other CRs the general approval of the study by the INT EC was enough.

On the 13th of April 2018, this last one was provided, while the local are going to be provided by the single collaborators.

The following table reports a summary of the present situation. At 31/10/2018, only three CRs have not yet received communication about approval by their local Ethical Committees (Bremen from Germany, Trento and Alto Adige from Italy), even if they are on proceedings.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Table 1: Cancer Registry in WASABY and Ethical Committee approvals.

Cancer Registry	Country	INT EC approval	Local approval
Belgium CR	Belgium	X	Retired from study
Bremen CR	Germany	X	Local EC: on proceedings
Schleswig-Holstein CR		X	INT EC approval is enough
Alto Adige CR	Italy	X	Local EC: on proceedings
Napoli CR		X	INT EC approval is enough
Palermo CR		X	INT EC approval is enough
Parma CR		X	INT EC approval is enough
Ragusa CR		X	INT EC approval is enough
Siracusa CR		X	INT EC approval is enough
Trento CR		X	Local EC: on proceedings
Umbria CR		X	Obtained by local EC
Varese CR		X	INT EC approval is enough
Lithuania CR	Lithuania	X	INT EC approval is enough
Greater Poland CR	Poland	X	Obtained by Polish National CR
Kielce CR		X	Obtained by Polish National CR
Silesia CR		X	Obtained by Polish National CR
Central Portugal CR	Portugal	X	Obtained by local Administration Board
Northern Portugal CR		X	Obtained by local EC
Slovenia CR	Slovenia	X	INT EC approval is enough
Basque Country CR	Spain	X	INT EC approval is enough
Castellon-Valencia CR		X	INT EC approval is enough
Girona CR		X	INT EC approval is enough
Granada CR		X	INT EC approval is enough
Murcia CR		X	INT EC approval is enough
Navarra CR		X	INT EC approval is enough
Northern Ireland	UK	X	INT EC approval is enough
French CRs (17)	France	X	Obtained collectively by referent

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 11 – M5.1 – LIST OF WP-5 EXPERTS

M6.1 Report

WASABY group of experts on spatial analysis

Version	*Author*	*Date*
V1	Ludivine Launoy & Elodie Guillaume	31.08.2018
V2	Ludivine Launoy & Elodie Guillaume	01.01.2019

Here is the list of expert in SES identified to develop EDI.

Country	**Contact**
Italy	Roberto Lillini, Marina Vercelli
Spain	Marc Saez, Marc Mari Dell' Olmo
Portugal	Ana Isabel Ribeiro
Slovenia	Vesna Zadnik
Germany	Ron Pritzkuleit
Poland	Krzysztof Czaderny
Nothern Ireland	Andy Moore, Bruna Pucci
Lithuania	

This list will be updated during the progress of the work of WP5. In December 2018 as Belgium leave the project, we update the previous table. We were contacted and meet Andy Moore and Bruna Pucci in December for Northern Ireland. Currently we still lacking a contact in Lithuania.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Wasaby report D5.1 on deprivation index available data

1. Introduction

The Wasaby project focuses on the geographical analysis of population-based cancer incidence data in connection with environmental factors, using breast cancer and water/soil contamination, as an example. In this project, the WP5 have to provide a national version of European Deprivation Index (or of other Deprivation Index) for all countries participating in the study, i.e., France, Germany, Italy, Lithuania, Poland, Portugal, Slovenia, Spain and Northern Ireland. In order to minimize the unavoidable ecological bias, the smallest area for which census data were available must be identified to include deprivation (assessed at aggregated level) as confounder in statistical modeling for breast cancer.

The main aim is to identify those participating countries for which a national version of the European Deprivation Index (EDI) is not available. For these countries, we have to collect information on census data availability, according to the data level, in order to estimate the possibility of a deprivation index construction. If EDI is not estimable, other deprivation indexes should be identified and collected. The present report focuses on the data available for estimating deprivation indexes in the various participating countries , and concludes on the feasibility of the European Deprivation Index construction in each country.

2. Background

2.1 Concept of deprivation

According to P. Townsend, deprivation refers to unmet needs caused by a general lack of resources, rather than financial needs alone, and needs vary between societies and periods. Since the single socioeconomic data are often absent or poorly collected in routine health databases, individual social status is regularly assessed through the socioeconomic characteristics of the place of residence, mainly at the small-area level of census blocks. Moreover, the aggregated-level index allows to consider for contextual factors not accounted by individual-level index. The purpose of the European Deprivation Index is to measure ecological deprivation at a small-area-level in a comparable manner across countries, despite the social and cultural specificities of the different countries, and the availability of census variables. The ecological deprivation indices are built according to shared methodological principles, by selecting fundamental needs associated with both objective and subjective poverty; they use the same theoretical concept of relative deprivation, through a European survey on relative deprivation (EU-SILC) that is regularly conducted on national samples across the EU. The concept of relative deprivation makes it possible to measure comparable social status using variables that may differ in each country.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

2.2 European deprivation index

The method for computing national versions of the European Deprivation Index (EDI) is described in different papers [Pornet C, JECH 2012 PMID: 22544918; Guillaume E, JECH, 2016 PMID: 26659762] and national versions of EDI are already available for 5 European countries (Italy, Portugal, Spain, England and France). This index is based on:

- the European survey on deprivation European Union Statistics on Income and Living Conditions (EU-SILC). EU-SILC is a cross-sectional and longitudinal sample survey providing data on income, poverty, social exclusion and living conditions in the European Union. From these data, the statistical office of the European Union (Eurostat—http://ec.europa.eu/eurostat/web/main) produces a European standardized questionnaire that is specifically designed to study deprivation. It consists of nine questions, common for all European Union members, aimed to evaluating needs that directly or indirectly induce financial inability. For each European Union member, the sum of weights for the sample design and the response rate to a national questionnaire were tailored on the basis of the national population size. All analyses were weighted for non-response and adjusted for sample design, so as to ensure the representativeness of the results for each member.
- The ecological data of the national population censuses Ecological data came from the last exhaustive national population censuses conducted in 2001 for Italy (Italian National Institute of Statistics: ISTAT), Portugal (National Institute of Statistics: INE), Spain (National Institute of Statistics: INE) and England (Office for National Statistics: ONS), and, in 1999, for France (National Institute for Statistics and Economic Studies: INSEE).

The method of EDI construction is based on three steps. The **first step** aims to construct an individual indicator for deprivation thanks to the identification of fundamental needs associated to objective and subjective poverty. The **second step** aims to identify variables available both at individual (EU-SILC survey) and aggregate levels (census) in each country. The **third step** is dedicated to the construction of ecological deprivation index. In this step, individual indicator of deprivation is explained in a logistic regression by variables identified in the previous step. The regression coefficients became the weights of the variables measured at aggregated level. The final index is the sum of these weighted variables.

3. Why and how will the deprivation index be applied in the WASABY Project

Since this study includes different European countries, the measurement of socioeconomic deprivation must be comparable or at least transferable between them despite the socio-cultural differences, so to improve comparability and reproducibility of the concept of relative deprivation. The European deprivation index is the only index with these characteristics.

EDI will be use in analyses of incidence and as a covariable to investigate the link between environmental pollutants and breast cancer incidence.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4. National and European Deprivation Index in the participating countries

4.1 EDI estimation

While EDI 2001 is already available in the aforementioned countries, the construction of EDI 2011 is in progress for all countries involved in WASABY. Table 1 here below describes the different steps of the construction progress in the Project.

Table 1: State of the art of EDI 2011 estimation for WASABY (at 31st March 2019)

	Year of Census	Status		
		Step 1	**Step 2**	**Step 3**
France	2011	X	X	X
Germany	2011	X		
Italy	2011	X	In progress	
Lithuania	2011	X		
Poland	2011	X		
Portugal	2011	X	X	X
Slovenia	2011	X	X	X
Spain	2011	X	In progress	
Northern Ireland	2011			

The French, Portuguese and Slovenian EDI-2011 were concluded. In a few weeks, Step 2 should be reached for Italy, and the same process will take place for Spain. At a later stage, Northern Ireland's EDI will be constructed, and so will the Polish and Lithuanian ones.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

4.2 National indices

For each country, the list of national available indices is not exhaustive, as it only concerns widely used indices.

France

The most frequently used index are the European Deprivation Index (EDI) and the French index of social deprivation (FDep). FDep is constructed by principal component analysis including four variables that are the percentage of: blue collars among active people, persons over 15 years of age with baccalaureate (upper secondary education degree), unemployed persons among active people, median income per household [REY G., RICAN S., JOUGLA E. (2011). Mesure des inégalités de mortalité par cause de décès - Approche écologique à l'aide d'un indice de désavantage social. BEH n°8-9 : pp 87-90]. FDep was initially constructed at municipality level but is now available at census block level.

Germany

According to scientific literature, the German deprivation index was built after the model of the British IMD.; most articles, however, are in German hence an accurate synthesis on the subject was not possible.

From an abstract in English, we know that a small-area, multidimensional Index of Multiple Deprivation (IMD) was developed for Germany, based on an established British method. Official sociodemographic, socioeconomic and environmental data was used to create a Bavarian Index of Multiple Deprivation (BIMD) (PMID: 22020751)

The German IMDs consist of seven deprivation domains representing single aspects of deprivation (income, employment and educational deprivation, municipal revenue deprivation, social capital deprivation, environment and security deprivation). Specific indicators were generated from data of official statistics, and assigned to the deprivation domains. The weighted single domains were finally combined to an overall index. The German IMDs are available at a municipal level and at a district level. (PMID: 29119206)

A version of the German Index of Multiple Deprivation GIMD based on data from 2007 to 2010 for all 412 rural and urban districts in Germany was then developed and mentioned in PMID: 25706042.

Italy

A nationwide deprivation index at municipality and census block level on Census data, from 2001 was developed and published in Italian. From the 280 variables defined at census block level (2001 General Census of Population and Housing) the following five traits that operationally combine to represent the multidimensionality of the social and material deprivation concept were selected: low level of education, unemployment, home non-ownership, single parent family and overcrowding. The index is calculated by summing standardized indicators. (PMID 21224518)

Lithuania

An index is not available neither at local nor at national level. Socioeconomic inequalities are studied using single indicators such as education (PMID 29408192, PMID 24227051), income, absolute material deprivation, assessed by the question "how often do you not have enough money for food" (PMID 24227051) , and relative deprivation defined by the number of household amenities (PMID 24227051).

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Poland

There are two local indices: the Poviat index [http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/6547/poviats_threatened_by_deprivation_2015.pdf] of deprivation and the local index of deprivation [http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/6543/local_concentration_of_deprivation_in_poland_2016.pdf] constructed at municipality level. These indices take into account population's income, unemployment, living conditions (flats with bathroom), education (results of lower-secondary school final exam) and access to goods and services (number of person per flat and percentage of children covered by the preschool education). These variables were summed up after normalization and divided by the number of variables included in the indices.

Portugal

In Portugal, no standard ecological deprivation index exists, contrasting with other countries. The Portuguese version of a transnational deprivation index, European Deprivation Index, was described in PMID: 28501033

Slovenia

No local deprivation measure is available. The Slovenian version of EDI was described in PMID: 29651315. EDI was recently developed at individual level (PMID: 30678244).

Spain

A few local indices are described (PMID: 25631857, PMID: 22846597) and no national index other than the EDI was constructed with 2001 census (PMID: 26659762). It is available at census tract level and must be updated with 2011 census.

Northern Ireland (UK)

The most common index are the Townsend and Carstairs indices and more recently the Index of Multiple Deprivation. The latter includes several dimension of deprivation such as income, employment, education, health, crime, access to housing services and living environment. For each one, variables that best reflect each dimension were taken into account [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464485/English_Indices_of_Deprivation_2015_-_Technical-Report.pdf]. All dimensions were then combined to obtain a measure of deprivation. An Index is available at small geographical unit-level (LSOA).

Availability of census data

A survey (cf Annex) was sent to each contact to establish the availability of data and the best geographical unit for the construction of EDI. Only one country did not answer to it (Lithuania). At the same time, national census institutes were contacted when more detail were considered to be useful.

A synthesis of the responses is presented in the table below describing the geographical units for which census data is available, and the number of inhabitants for all of them. Units are presented from the smallest geographical units (level 1) to the largest ones (table 2).

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Table 2: Geographical units for each country involved in WASABY

	Smallest units (level 1)			Units just higher than the smallest (level 2)			Units level 3			Units level 4			Census
	Name	**Number of unit**	**Average population (min-max)**	**Name**	**Number of unit**	**Average population (min-max)**	**Name**	**Number of unit**	**Average Population (min-max)**	**Name**	**Number of unit**	**Average population (min-max)**	
France	IRIS	50,867	1,277 (0-11,159)	Municipality	36,664	1,443 (0-1,926,595)	Department			Region			
Germany	Municipality	11,054	1,719 (9-3,469,849)										
Italy	Census section	366,863	165 (0-7,647)	Municipality	8,092	7,492 (30-2,617,000)	Province	110	551,149 (57,657-4,094,659)	Region	21	2,886,973 (127,425-9,809,298)	Exhaustive, Traditional, 10 years
Lithuania	Census block	71	Not yet available	Seniunijos	546	Not yet available	Municipalities	60	Not yet available	Apskritis	10	Not yet available	
Poland	Municipality (Gmina)	2,478	15,510 (1,302-1,764,615)	Poviat	380	101,141 (20,270-1,764,615)	Subregion	73	526,487 (189,469-1,764,615)	Voivodeship	16	2,402,097 (990,069-5,384,617)	Sample-based survey
Portugal	Census block	265,955	40 (0-1,742)	Census block group	18,074	584 (0-2,152)	Parish	4,260	2,479 (31-66,250)	Municipality	308	34,293 (1,430-547,733)	Exhaustive, Traditional, 10 years
Slovenia	Settlements	5,972	338 (1-259,896)	Polling stations for National Assembly elections	3,104	660 (30-4,516)	Municipalities	210	9,762 (316-280,140)	Administrative units	58	35,348 (8,416-347,147)	Exhaustive, Registered-based census, 3 to 4 years
Spain	Census tracts	525	3,384 (85-95,675)	Municipalities	221	3,384 (95-95,675)							
Nothern Ireland	Census Output Area	5,022	350 (100-2,100)	Small Areas	4,537	400 (98-3,072)	Super Output Areas	890	2,000 (900-4,200)	Electoral Wards	582	3,000 (700-9,500)	Exhaustive, Traditional, 10 years

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

For countries like France, Portugal, Spain (even if the largest population is higher in this country), Italy, Northern Ireland and Slovenia, geographical units are in accordance with studying social inequalities. This unit must be the smallest available so as to limit ecological fallacy (Woods et al, 2006). For these countries, average population for each spatial unit are equal to 40 for Portugal, 165 for Italy, 350 for Northern Ireland, 1,277 for France and 3,384 for Spain.

For countries like Poland, the only geographical units available are those at municipality-level. Average number of inhabitants is equal to 15,510 for Poland. We expect the socioeconomic inhabitant composition in such units to be very heterogeneous but there is no alternative if census data at a smallest unit- level is missing.

For Germany, the number of inhabitants in the different municipalities is highly inhomogeneous from one federal state to another. Moreover, socioeconomic data will be obtained through different sources and may not be available for all federal states. For this reason, the construction of EDI does not seem to be relevant for this country (table 2), hence the use of a national index is recommended.
The geographical unit level used for the development of EDI is described in table 3, below. This also ensures that we proceed in line with the population data collected within WASABY WP4.

Table 3: Geographical units for EDI construction

Country	Census block	Parish level	Municipality level
France	*		
Germany		NOT RELEVANT	
Italy	*		*
Lithuania			*
Poland			*
Portugal		*	*
Slovenia	*		
Spain	*		*
Northern Ireland			*

The contact names for the development of EDI in each country are entered below:

Country	Contact	Institute
France	Ludivine Launay	Caen University
	Elodie Guillaume	Caen University
Germany	Ron Pritzkuleit	Lübeck University
Italy	Roberto Lillini	Instituto Nazionale Tumori (INT)
Lithuania	Ieva Vincerzevskiene	Lithuanian Cancer Registry
Poland	Krzysztof Czaderny	Centrum Onkologii – Instytut im. Marii Skłodowskiej-Curie
Portugal	Ana Isabel Correia Ribeiro	EPIUnit - Public Health Institute University Porto
Slovenia	Vesna Zadnik	Oncology institute in Ljubljana
Spain	Marc Saez	Girona University
	Marc Mari Dell'Olmo	Agència de Salut Pública de Barcelona
Northern Ireland	Bruna Pucci	Ulster University
	Adrian Moore	Ulster University

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Conclusions

Except for Lithuania for which we need a further contact to progress in our work, we have compiled the information concerning the census level data and we have identified a country for which EDI cannot be developed (Germany).

We have the information on the different level for which census data exists but we must remain vigilant about the actual availability of census data. We want to distinguish the fact that we have information about census level but we are not sure to have the access of these data for all countries. For example, in Slovenia, census data were only available on the census institute laptop.

Our work shows that in some countries (Germany, Poland) municipalities are the census data smallest unit. In such cases, we could technically develop an index at this level but we prefer to be careful about the relevance of using EDI on this level for this project. Indeed making the assumption that all individuals living in the same geographical units have the same socioeconomic level is a very high assumption. According to the ecological fallacy, it is all the more true that the unit is large (as the municipality). Moreover, heterogeneity within a unit such as the municipality is even more important in rural areas than in urban areas.

At this stage, the indicators for estimating the European Deprivation index in WASABY are:

- Number of countries for which we are sure about the EDI 2011 estimation: 5 (France, Spain, Italy, Portugal, Slovenia)
- Number of countries for which we envisage the EDI 2011 estimation: 3 (Lithuania, Poland, Northern Ireland)
- Number of countries for which we will not be able to do EDI estimation : 1 (Germany)
- Number of countries estimating EDI for the first time: 3 (Lithuania, Poland, Northern Ireland)

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Annex

WASABY - Description report on geographical area in the country

Country: _____

Contact name: _____

As you know, your country is involved in WASABY Project which aim to identify areas with higher cancer rates, so to study whether pollutant contamination may be a cause for increased cancer risk. As part of Work Package 5 (construction of the European Deprivation Index, EDI), we need some information to determine the best geographical unit to compute EDI for your country. Indeed, to limit ecological bias, EDI have to be computed at the smallest geographical area for which census data are available. This area will determine the geocoding level for registries data for each country. That is why we sent you this survey in addition to the one sent by WP4. Do not hesitate to contact people who can answer to these questions in your country or to indicate the more pertinent contact to obtain these informations.

About geographical area

For some countries, socioeconomic data could be not available in national census data but available in regional statistics as in Germany. So we need to investigate the availability in census of such information for the different geographical unit.

Can you cite the different geographical area (administrative and non-administrative) in your country (for example, district, municipality, census block...). Cite them by the smallest to the largest scale. For all of them, precise if census data are available. Provide a little description if some area are not disposable for all units (for example in Germany, administrative region does not divide all Bundesländer). If some geographical area are dependent of the number of population (or another variable), please precise it in description (for example in France, IRIS concerned only all the municipalities with more than 10,000 people, a part of the municipalities with more than 5,000 inhabitants and less than 10,000; municipalities with less than 5,000 inhabitants are not divided and are an IRIS by themselves).

For example, in France:

Geographical area	Description	Number of units	Mean	min	max	Corresponding NUTS	Availability of the census
Region						2	Yes
Department						3	Yes
Municipality		36,664	1,443	0	1, 926,595	LAU	Yes
IRIS	All municipalities with more than 10,000 inhabitants, a part of municipalities with more than 5,000 inhabitant and less than10,000; municipalities with less than 5,000 inhabitants are not divided and are an IRIS by themselves	50,867	1,277	0	11,159		Yes
Grid							Little information

1

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

WASABY - Description report on geographical area in the country

			Population					
Geographical area	Description (if necessary)	Number of units (if available)	Mean (if available)	Min (if available)	Max (if available)	Corresponding NUTS* (if available)	Availability of the census	Socioeconomic indicator available

*The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU: NUTS 1=major socio-economic regions, NUTS 2 = basic regions for the application of regional policies, NUTS 3: small regions for specific diagnoses. To meet the demand for statistics at a local level, Eurostat maintains a system of Local Administrative Units (LAUs) compatible with NUTS. These LAUs are the building blocks of the NUTS, and comprise the municipalities and communes of the European Union. (http://ec.europa.eu/eurostat/web/nuts/background; http://ec.europa.eu/eurostat/web/nuts/local-administrative-units).

2

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 2017-2020 under Grant Agreement PP-2-5-2016 (# 769767)

WASABY

WASABY - Description report on geographical area in the country

About census

What is the name of the institute in charge of it: _____

What is its website: _____

Is the census exhaustive?

- ☐ Yes
- ☐ No, Precise the size of the survey sample: _____

What is the type of census?

- ☐ Traditional
- ☐ Rolling
- ☐ Sample based administrative data based
- ☐ Sample-based
- ☐ Other (Precise:_____)

Is it conducted annually?

- ☐ Yes
- ☐ No. Please provide the specific timing:_____

Does it include socioeconomic data?

- ☐ Yes
- ☐ No. Precise, if these data are available by another mean and how: _____

Reference

If there is some reference that can complete these information, please, write it:

1. _____
2. _____
3. _____
4. _____

3

WASABY

WASABY - Description report on geographical area in the country

Contacts:

Elodie GUILLAUME

elodie.guillaume@unicaen.fr

Ludivine LAUNAY

ludivine.launay@inserm.fr

4

ANNEX 13 – M6.1 – LIST OF WP-6 EXPERTS

M6.1 Report

WASABY group of experts on spatial analysis

Version	*Author*	*Date*
V1	Tina Žagar	18.07.2018

Background

In the framework of the WASABY Project (http://www.wasabysite.it/) the Work Package 6 (WP-6) is aimed to determine methodologies suitable to perform spatial analysis with cancer data, using breast cancer incidence as an example. In order to achieve this aim we established a group of experts in spatial analysis techniques able to be performed with routinely data collected by cancer registries. The WP-6 has one milestones to be reached in months 3 to 7 according to Gantt chart: M6.1 – Creation of WP-6 group of experts on spatial analysis.

The group of experts on spatial analysis will determine the methods suitable for risk analysis of cancer incidence provided by cancer registries according to availability of data at different spatial aggregation level with emphasis on small-area level and identify open source software(s) applicable in these methods. Furthermore, the group will prepare a practical handbook (including details on data preparation) for cancer registries' personnel and other researchers aiming to promote spatial analysis of cancer registries' data in general.

The group of experts on spatial analysis

The experts on spatial analysis were identified through already established work and also according to the participation in the WASABY project or other past international collaborating actions on spatial analysing of the cancer registries' data.

The members of WASABY group of experts on spatial analysis (Table 1) are possible to change later on during the course of the project – additional members could be invited into the group or existing could be removed from the list.

Table 1: Members of WASABY group of experts on spatial analysis.

Name	Affiliation	Country	e-mail adress
Tina Žagar	Slovenian Cancer Registry Inctitute of Oncology Ljubljana Zaloška 2 1000 Ljubljana	Slovenia	tzagar@onko-i.si
Vesna Zadnik	Slovenian Cancer Registry Inctitute of Oncology Ljubljana Zaloška 2 1000 Ljubljana	Slovenia	vzadnik@onko-i.si
Andreja Kukec	Faculty of Medicine University of Ljubljana Vrazov trg 2 1000 Ljubljana	Slovenia	andreja.kukec@mf.uni-lj.si
Ron Pritzkuleit	Institute for Cancer Epidemiology Universitat zu Lubeck Ratzeburger Allee 160 23538 Lubeck	Germany	ron.pritzkuleit@uksh.de
Marc Colonna	Isere Cancer Registry CHU de Grenoble – Pavillon E BP 217 38043 Grenoble Cedex 9	France	mcolonna.registre@wanadoo.fr
Joséphine Bryere	Universite de Caen Normandie 3 Avenue Général Harris 14076 Caen Cedex 05	France	josephine.bryere@gmail.com
Ludivine Launay	Universite de Caen Normandie 3 Avenue Général Harris 14076 Caen Cedex 05	France	ludivine.launay@inserm.fr
Fortunato Bianconi	Umbria Cancer Registry University of Perugia Via del Giochetto 06126 Perugia	Italy	fortunato.bianconi@gmail.com
Maurizio Zarcone	Palermo Cancer Registry Palermo University Hospital Via del Vespro 133 90127 Palermo	Italy	registrotumoripalermo@unipa.it
Martina Bertoldi	Varese Province Cancer Registry Fondazione “Istituto Nazionale Tumori” Via Venezian 1 20133 Milan	Italy	martina.bertoldi@istitutotumori.mi.it
Eero Pukkala	Finnish Cancer Registry Institute for Statistical and Epidemiological Cancer Research Unioninkatu 22 FI-00130 Helsinki	Finland	eero.pukkala@cancer.fi

ANNEX 14 – REPORT WITH ANALYSIS OF VARESE CR

From Varese CR data to analysis: first results

After the 1st Steering Committee, the Varese CR provided the first complete set of data, geo-coded at Census tract level.

WASABY's WP4 group chose to test some spatial analysis techniques on their dataset:

- SARAR models by STATA 14 (module: spreg)
- CARBAYES models by R (implementing BYM models).

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step by step discussion on the best model and software to be used for WASABY's spatial analysis, at Census tract level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 1996 to the 31st December 2012, for a total number of 2679 female breast cancer patients aged 0-49 year.

Merging observed 0-49 cases by Census tract with Varese province shapefile and EDI 2001 dataset

A dataset with CTs as key variable was built, merging the information from the Varese province shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2001 dataset with the geo-coded Varese CR dataset.

The result was a dataset in which the rows reported the CTs with the sum of the observed cases (from the Varese CR) and all the other information at CT level (from the other datasets).

The 2679 incident cases were distributed in 1457 CTs. In Census tracts (CTs) where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b).

Result	# of obs.
not matched CTs	401
matched CTs	4,793
total CTs	5,194

Not-matching 401 CTs presented missing values in Census data and, consequently, in SIRs. These were necessary for mapping the area completely and avoid biases in the SARAR models application. Hence, missing values and SIR were recoded to 0.

After this action, all 5194 CTs of the Varese province were allowed to be used in the SARAR models.

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Italian (`sir049_ita`), European (`sir049_eu`) and World (`sir049_wor`) population standard were used.

The following table reports the average SIRs for the Varese province, along with the standard error and the 95% C.I.

	Mean	Std. Err.	[95% Conf.	Interval]
sir049_ita	1.724784	.1257418	1.478277	1.97129
sir049_eu	1.811253	.1320457	1.552388	2.070118
sir049_wor	2.367196	.1725755	2.028875	2.705517

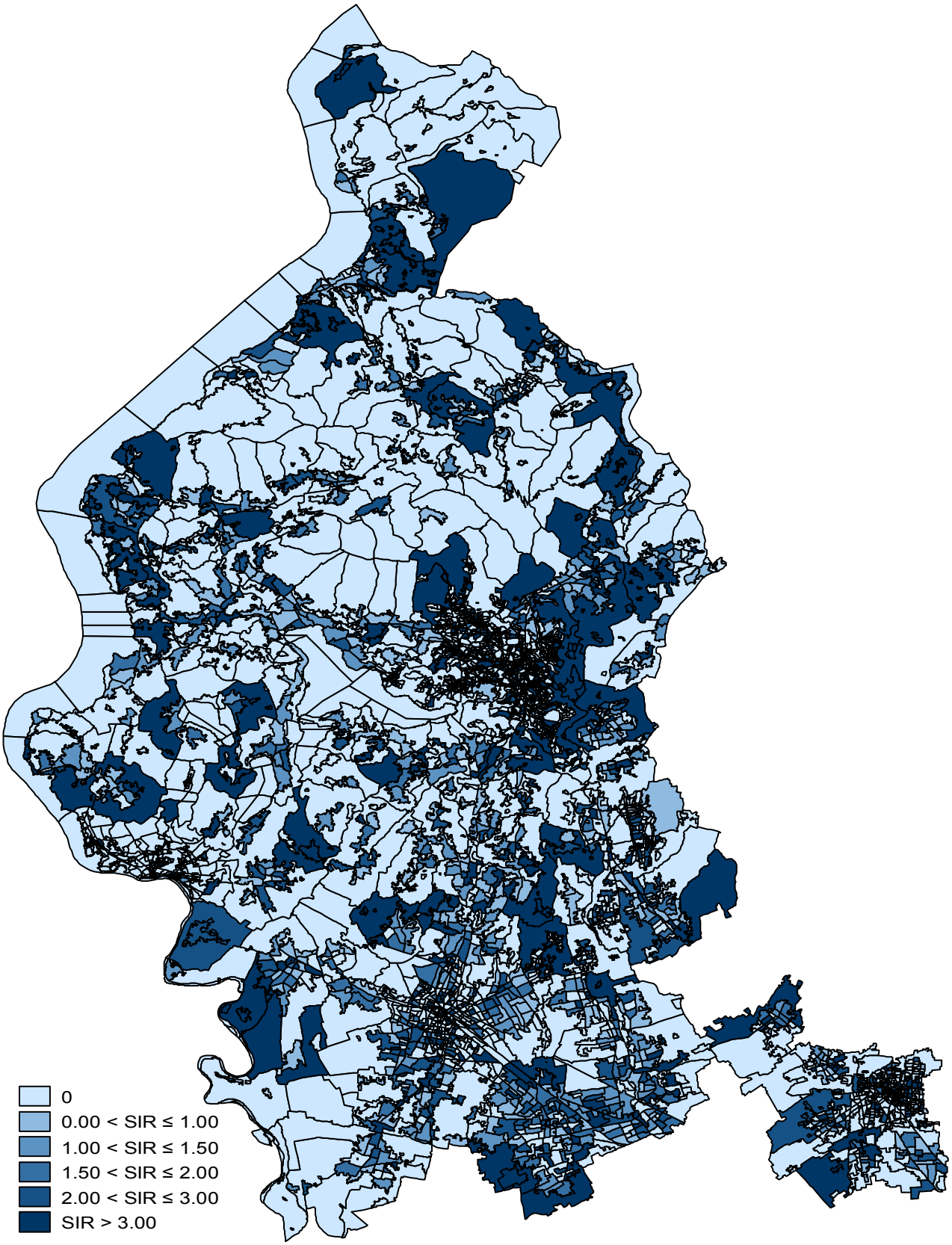
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
$0.00 < SIR \leq 1.00$
$1.00 < SIR \leq 1.50$
$1.50 < SIR \leq 2.00$
$2.00 < SIR \leq 3.00$
$SIR > 3.00$

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by European standardization was reported. No smoothing was applied.

Standard: Europe



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object varemat
------------------------------------------
        Matrix |                Description
---------------+--------------------------
    Dimensions |                5194 x 5194
     Stored as |                5194 x 5194
 Links         |
         total |                      30646
           min |                          1
          mean |                    5.90027
           max |                        100
------------------------------------------
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by EDI 2001). The EDI 2001 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Italy

```
Spatial autoregressive model                    Number of obs   =      5194
(GS2SLS estimates)

------------------------------------------------------------------------------
     sir049_ita |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049_ita      |
zediweighted_01 |  -.1241949   .0540681    -2.30   0.022    -.2301664   -.0182234
          _cons |   .9481836   .2090618     4.54   0.000     .5384301    1.357937
----------------+-------------------------------------------------------------
lambda          |
          _cons |   6.301074   1.297145     4.86   0.000     3.758716    8.843432
----------------+-------------------------------------------------------------
rho             |
          _cons |  -12.38115   1.510662    -8.20   0.000    -15.34199   -9.420304
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =      5194
(GS2SLS estimates)

------------------------------------------------------------------------------
      sir049_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049_eu       |
zediweighted_01 |  -.1304213   .0567787    -2.30   0.022    -.2417055    -.019137
          _cons |   .9957195   .2195428     4.54   0.000     .5654235    1.426015
----------------+-------------------------------------------------------------
lambda          |
          _cons |   6.301074   1.297145     4.86   0.000     3.758716    8.843432
----------------+-------------------------------------------------------------
rho             |
          _cons |  -12.38115   1.510662    -8.20   0.000    -15.34199   -9.420304
------------------------------------------------------------------------------
```

Standardization: World

```
Spatial autoregressive model                      Number of obs   =      5194
(GS2SLS estimates)

------------------------------------------------------------------------------
     sir049_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049_wor      |
zediweighted_01 |  -.1704525   .0742063    -2.30   0.022    -.3158942   -.0250109
          _cons |   1.301344   .2869289     4.54   0.000     .7389736    1.863714
----------------+-------------------------------------------------------------
lambda          |
          _cons |   6.301074   1.297145     4.86   0.000     3.758716    8.843432
----------------+-------------------------------------------------------------
rho             |
          _cons |  -12.38115   1.510662    -8.20   0.000    -15.34199   -9.420304
------------------------------------------------------------------------------
```

The graphical representation was reported only for the European standardization.

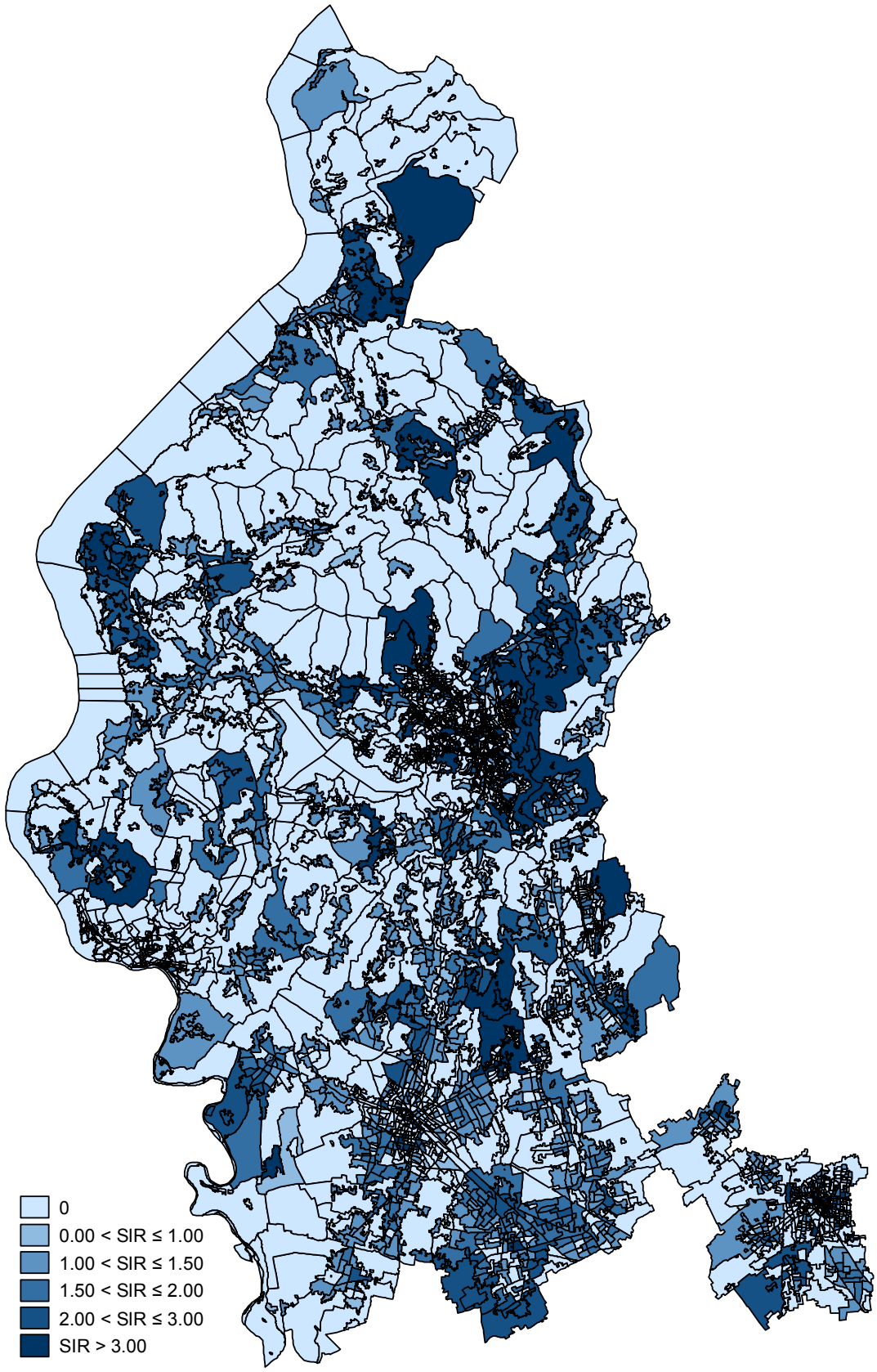
Moreover, also a zoom on the Varese municipality, the largest and most populous of the province, was reported.

References

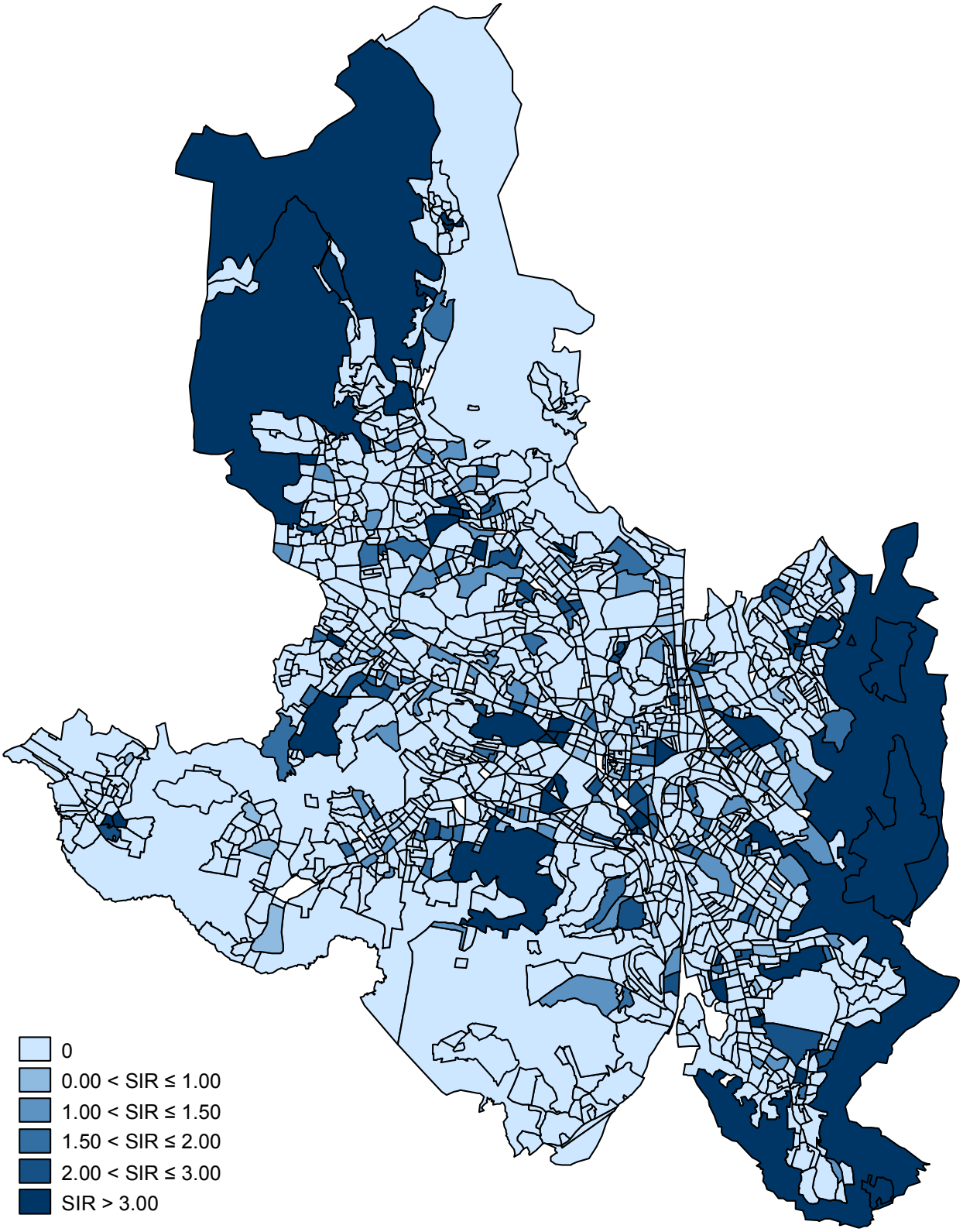
Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

Standardization: Europe



Zooming on Varese municipality:



ANNEX 15 – REPORT WITH ANALYSIS OF RAGUSA CANCER REGISTRY DATA

From Ragusa CR data to analysis: first results

Between September and October, the Ragusa CR provided the complete set of data, geo-coded at Census Tract (SEZ2001) level, integrating some correction about the right attribution of the 2001 SEZ2001.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step-by-step discussion about the best model and software to be used for WASABY's spatial analysis, at Census Tract (named SEZ2001 for Ragusa) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2001 to the 31st December 2012, for a total number of 515 female breast cancer patients aged 0-49 year.

The Italian National Statistics Office provided the 2001 shapefile at SU level for the registry-covered area, the Ragusa CR provided the 2001 Census Population and the INT provided the 2001 European Deprivation Index (EDI) at SEZ2001 level.

Merging observed 0-49 cases by Census tract with Ragusa province shapefile and 2001 EDI dataset

A dataset with SEZ2001 as key variable was built, merging the information from the Ragusa shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2001 dataset with the geo-coded Ragusa CR dataset.

The result was a dataset in which the rows reported the SEZ2001 with the sum of the observed cases (from the Ragusa CR) and all the other information at SEZ2001level (from the other datasets).

Only 499 cases reported the correct address and information for geo-coding; 15 cases were excluded because no attribution of SEZ2001 was available.
The 499 considered incident cases were distributed in 371 SEZ2001s. In 4046 SEZ2001s where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Italian (`sir049f_ita`), European (`sir049f_eu`) and World (`sir049f_wor`) population standard were used.

The following table reports the average SIRs for the Ragusa province, along with the standard error and the 95% C.I.

```
Mean estimation                   Number of obs    =      4,417

------------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+----------------------------------------------------
 sir049f_ita |   1.023566   .1124349       .803137    1.243995
  sir049f_eu |   1.194678    .131231      .9373994    1.451956
 sir049f_wor |   1.766777    .194074      1.386295    2.147259
------------------------------------------------------------------
```

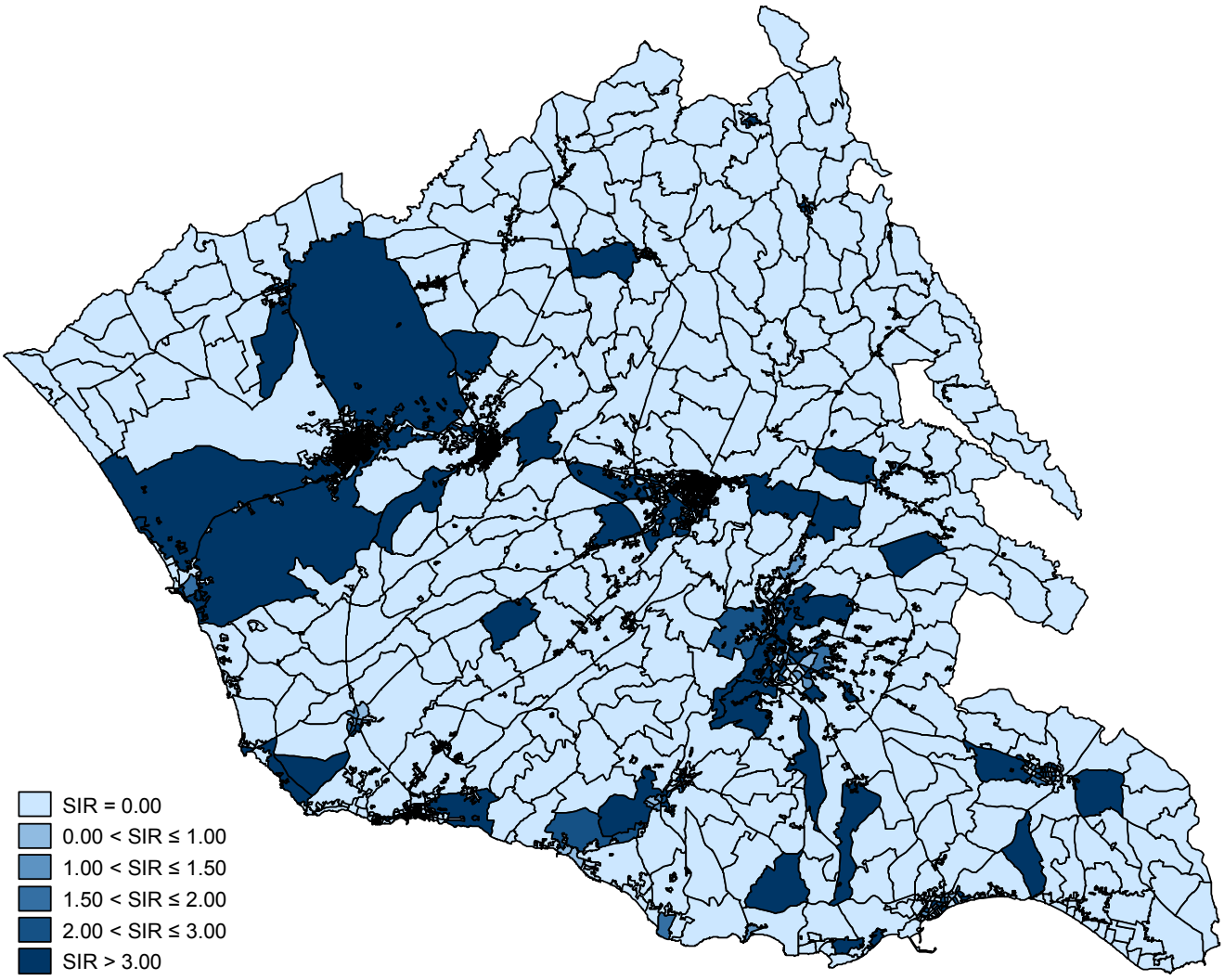
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by World standardization was reported. No smoothing was applied.

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object ragumat
------------------------------------------
        Matrix |                Description
---------------+--------------------------
    Dimensions |                4417 x 4417
     Stored as |                4417 x 4417
 Links         |
         total |                      29076
           min |                          0
          mean |                   6.582748
           max |                        103
------------------------------------------
warning: spatial-weighting matrix contains 1 island
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by DE 2011). The EDI 2001 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Italy

```
Spatial autoregressive model                    Number of obs   =      4417
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_ita |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_ita     |
zediweighted_01 |  -.1108398   .0376302    -2.95   0.003    -.1845936   -.0370859
          _cons |   1.331692   .4641391     2.87   0.004     .4219963    2.241388
----------------+-------------------------------------------------------------
lambda          |
          _cons |  -14.83249   1.043814   -14.21   0.000    -16.87833   -12.78665
----------------+-------------------------------------------------------------
rho             |
          _cons |   12.43581   .5698921    21.82   0.000     11.31885    13.55278
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =      4417
(GS2SLS estimates)

------------------------------------------------------------------------------
     sir049f_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_eu      |
zediweighted_01 |  -.1293691   .0439209    -2.95   0.003    -.2154526   -.0432857
          _cons |   1.554315   .5417304     2.87   0.004     .4925425    2.616087
----------------+-------------------------------------------------------------
lambda          |
          _cons |  -14.83249   1.043814   -14.21   0.000    -16.87833   -12.78665
----------------+-------------------------------------------------------------
rho             |
          _cons |   12.43581   .5698921    21.82   0.000     11.31885    13.55278
------------------------------------------------------------------------------
```

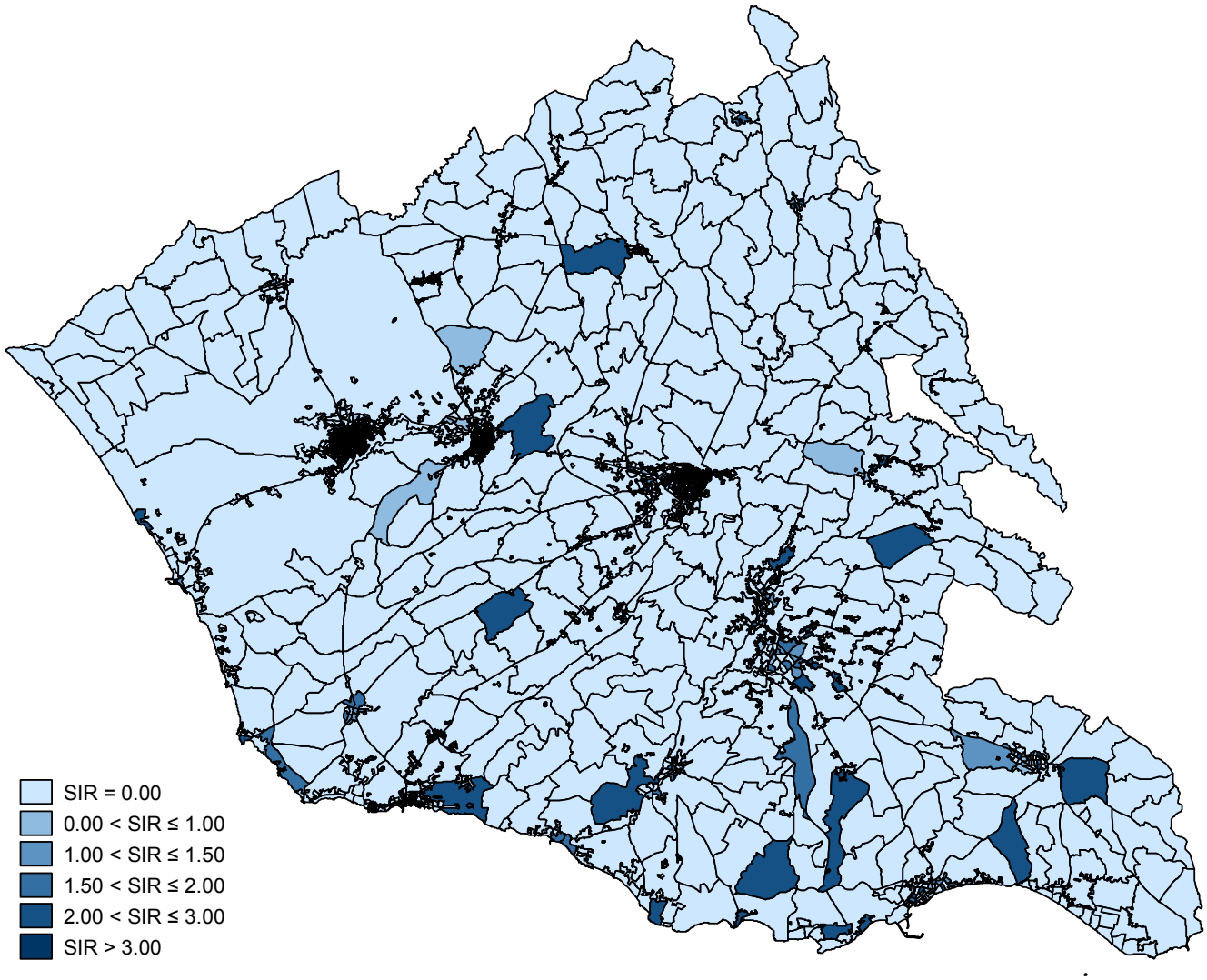
Standardization: World

```
Spatial autoregressive model                    Number of obs   =      4417
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_wor     |
zediweighted_01 |  -.1913205   .0649535    -2.95   0.003    -.3186271    -.064014
          _cons |   2.298634   .8011507     2.87   0.004     .7284078    3.868861
----------------+-------------------------------------------------------------
lambda          |
          _cons |  -14.83249   1.043814   -14.21   0.000    -16.87833   -12.78665
----------------+-------------------------------------------------------------
rho             |
          _cons |   12.43581   .5698921    21.82   0.000     11.31885    13.55278
------------------------------------------------------------------------------
```

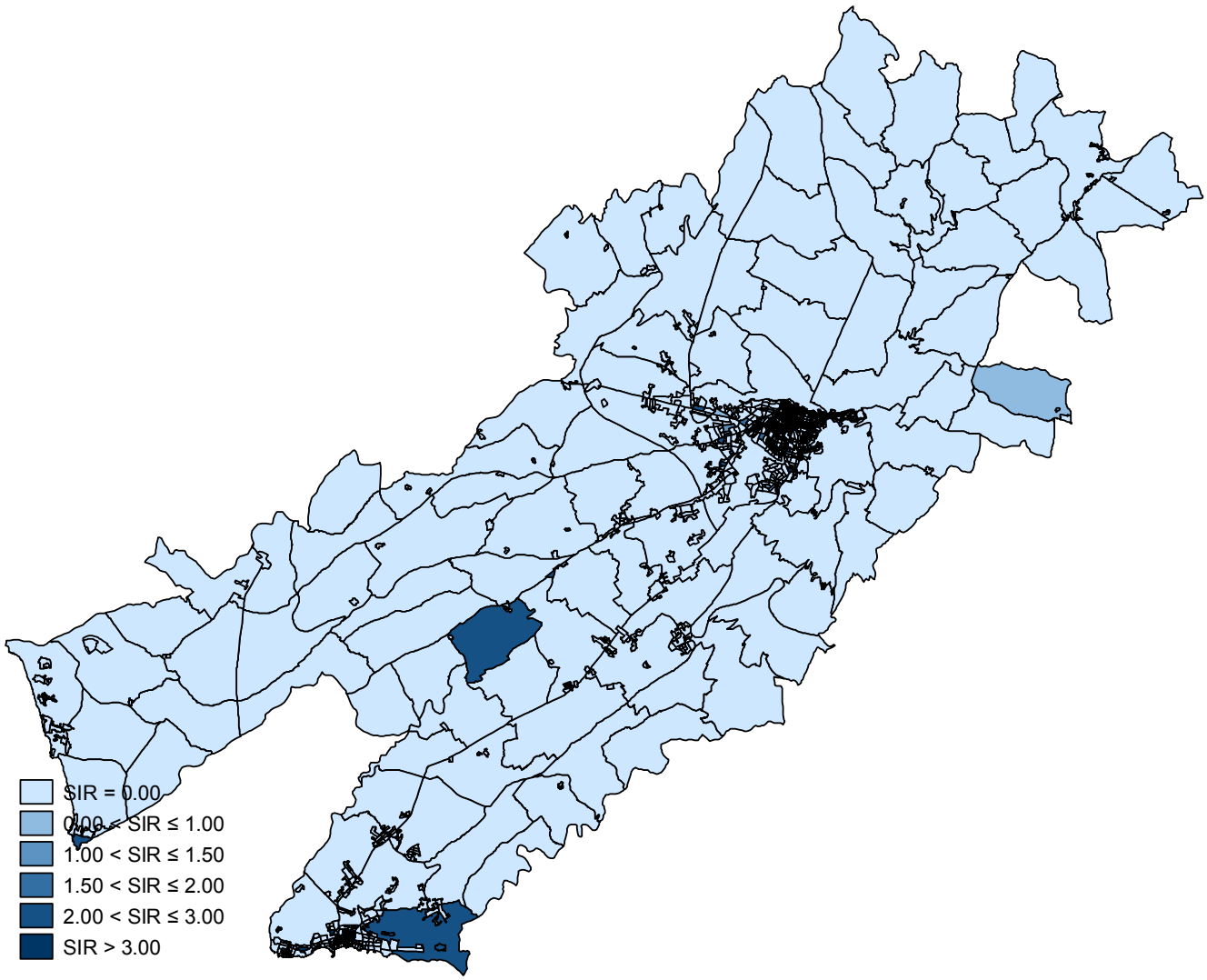
The graphical representation was reported only for the World standardization.

Moreover, also a zoom on the Ragusa, Vittoria, Modica, Comiso and Scicli municipalities, the largest and most populous ones of the province, were reported.

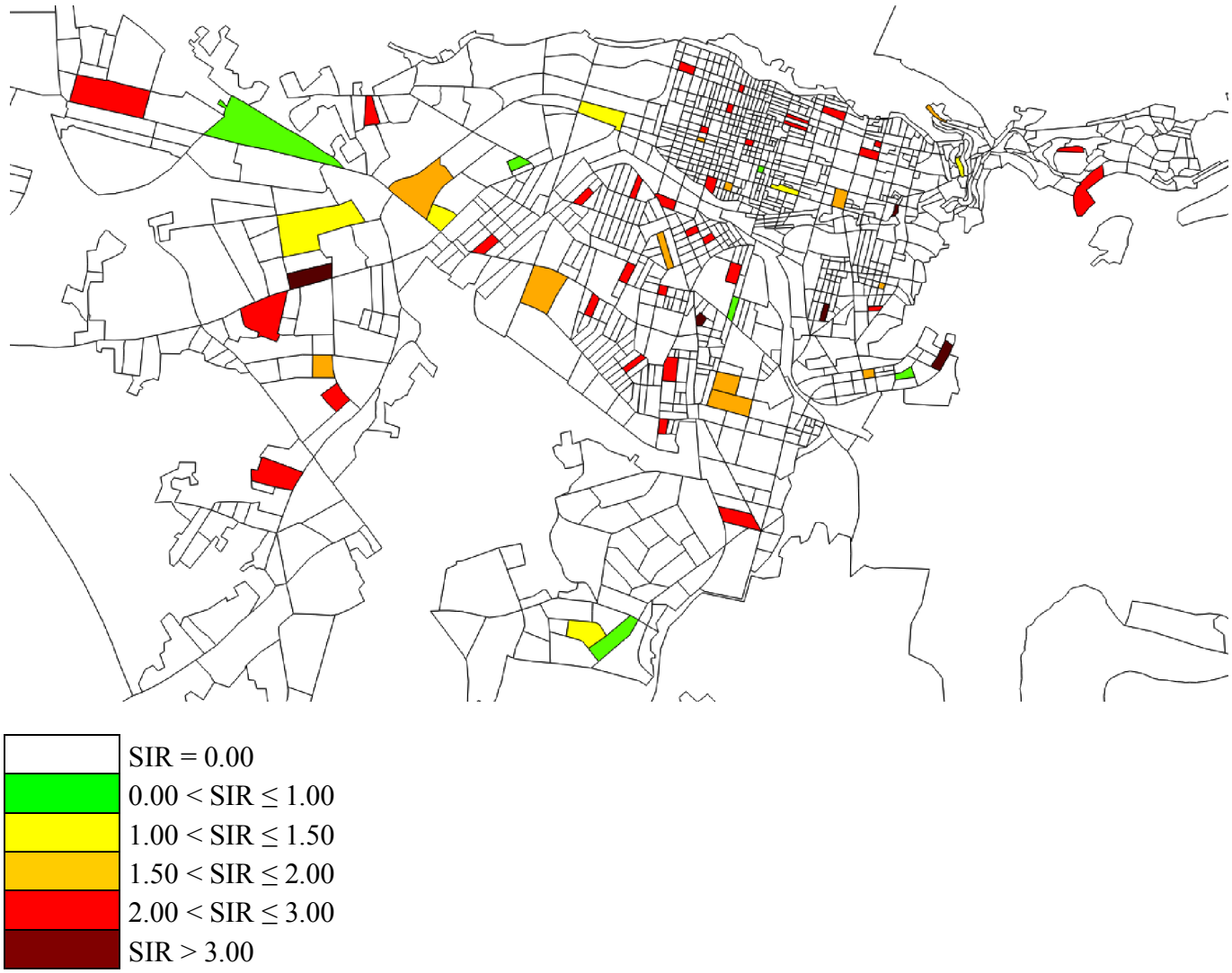
Standardization: World



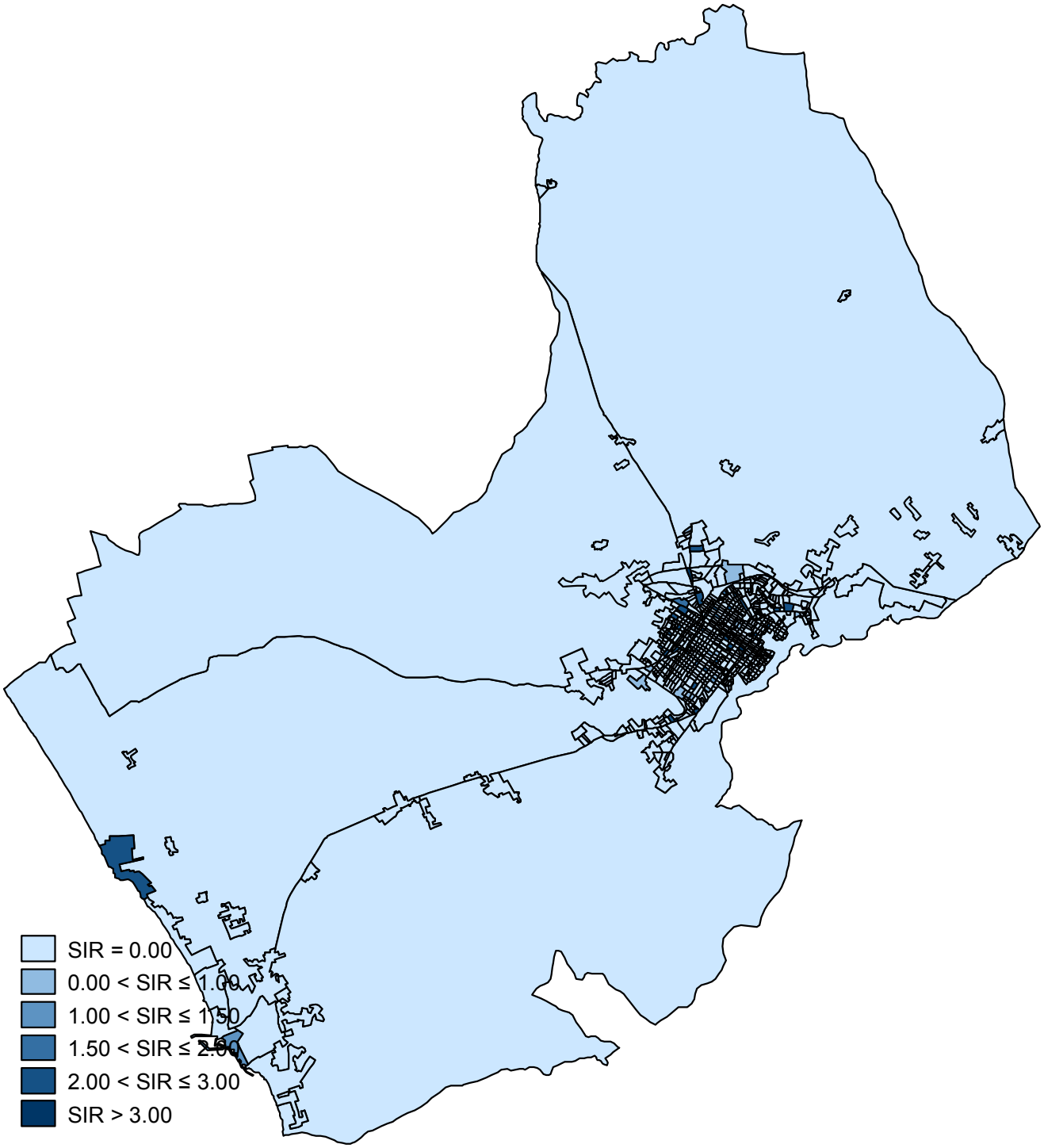
Zooming on Ragusa municipality:



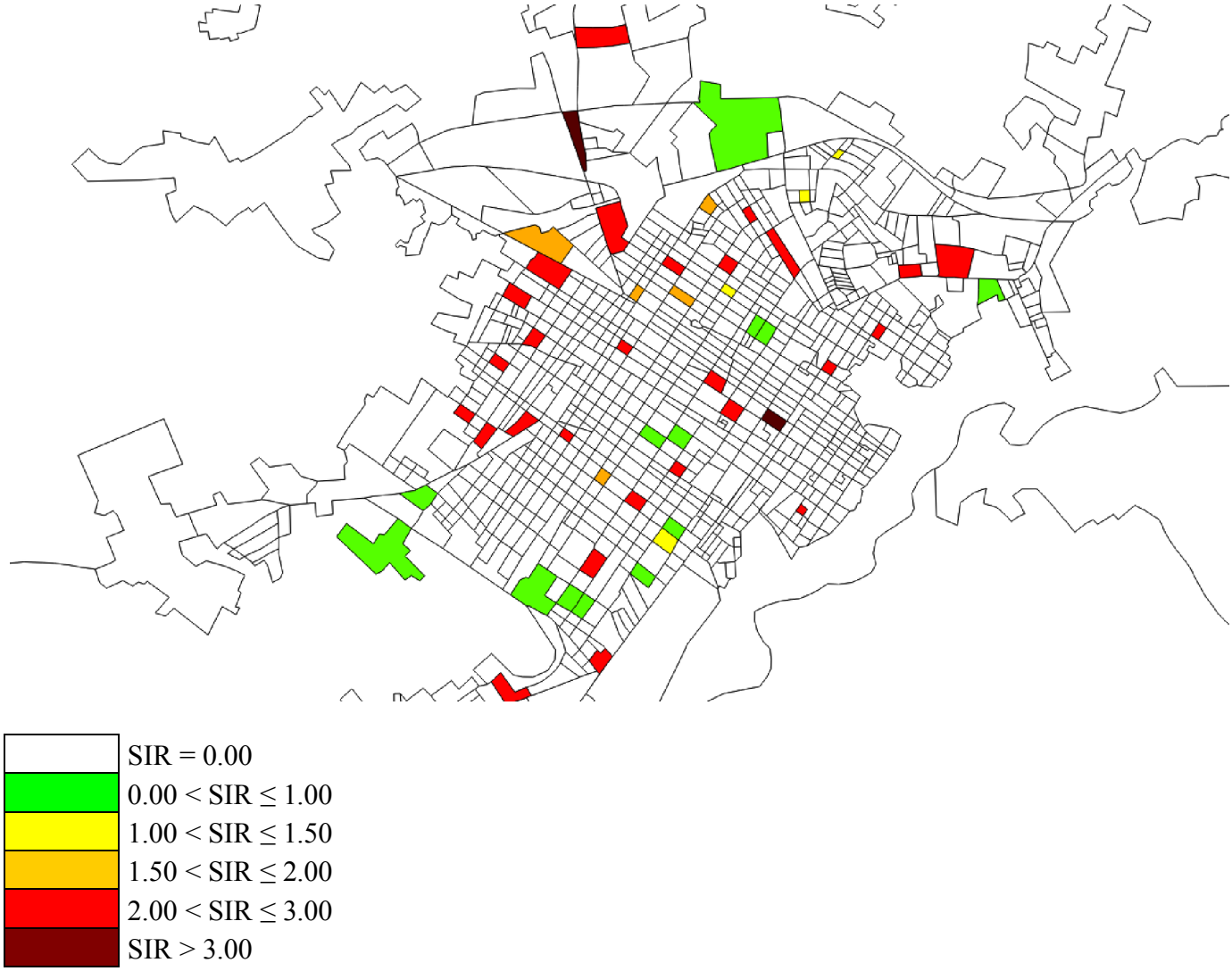
Zooming on Ragusa municipality centre:



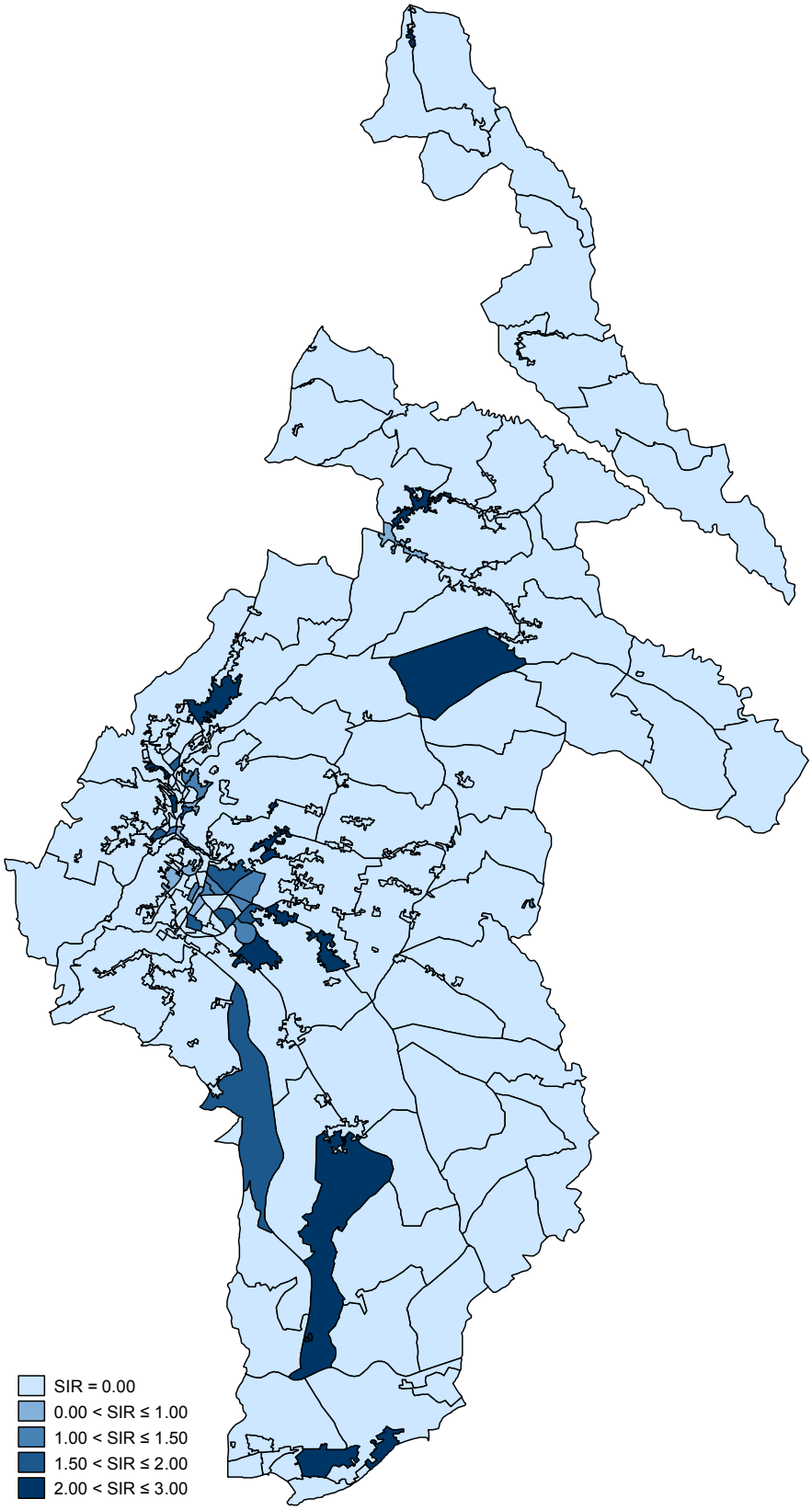
Zooming on Vittoria municipality:



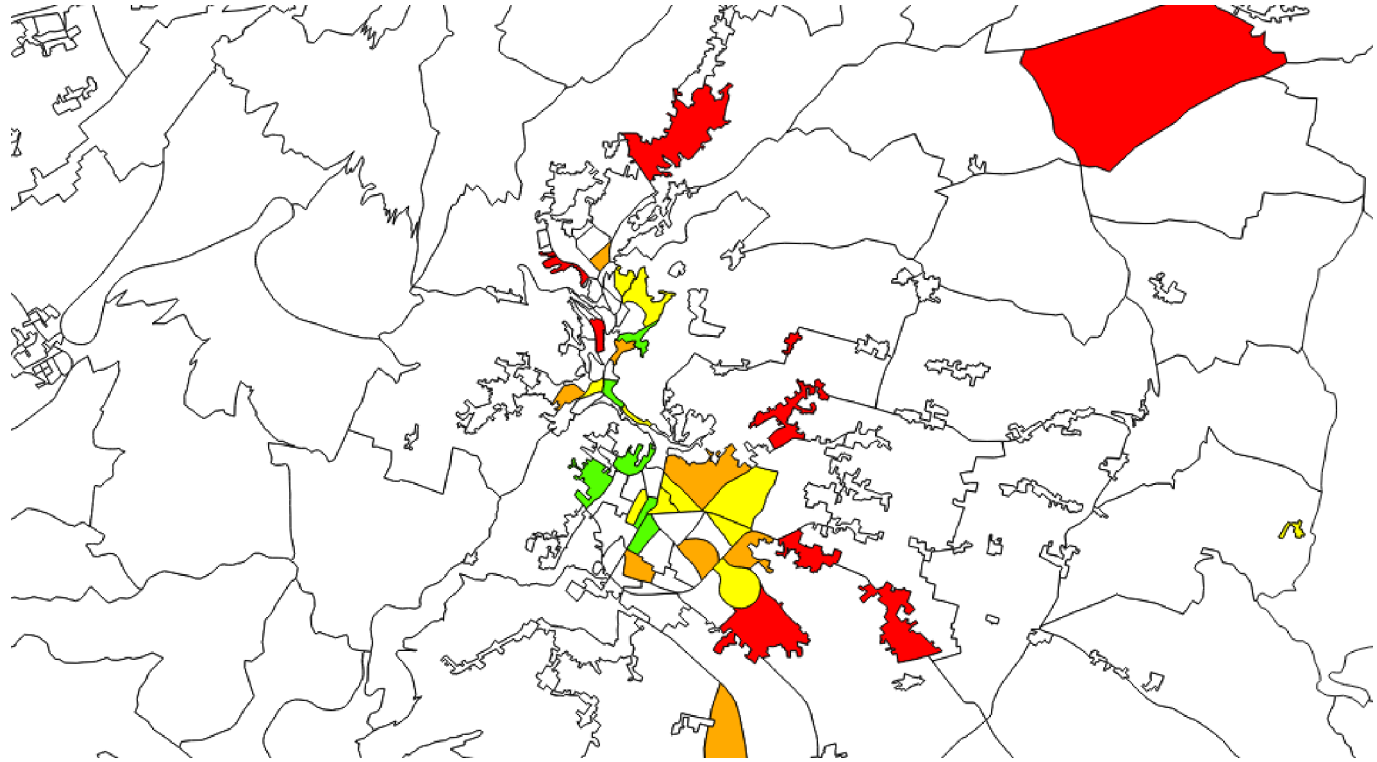
Zooming on Vittoria municipality centre:



Zooming on Modica municipality:

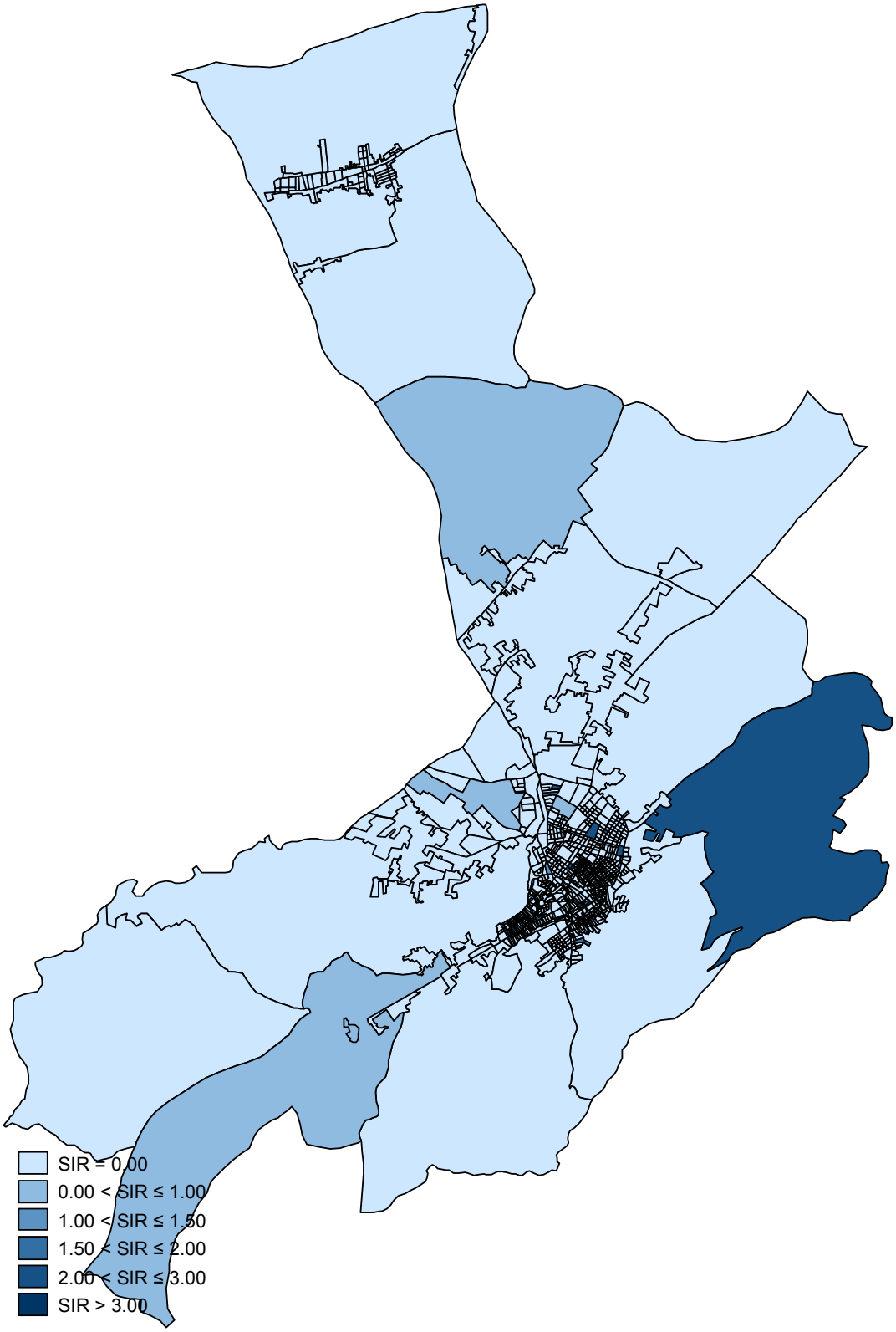


Zooming on Modica municipality centre:

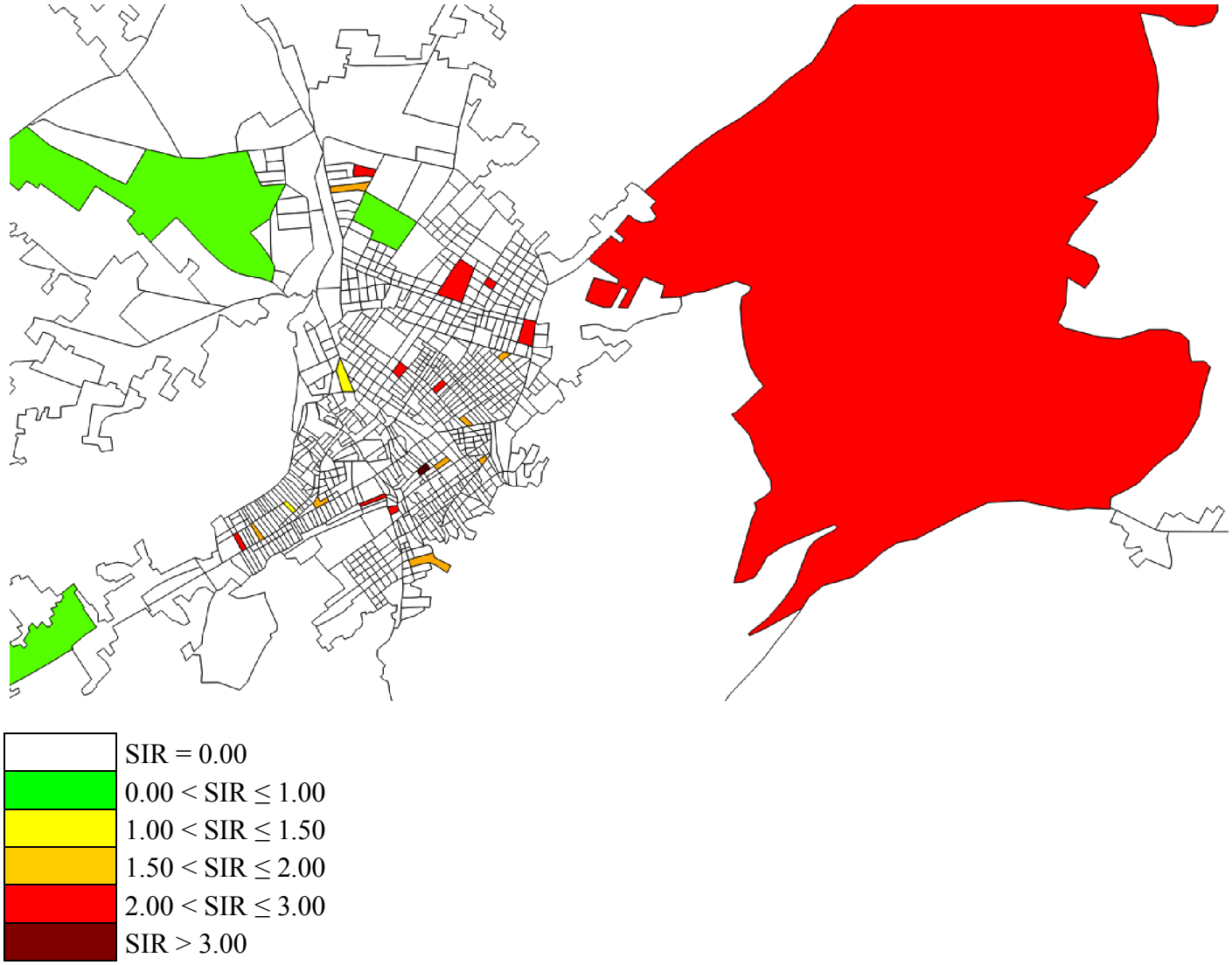


SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

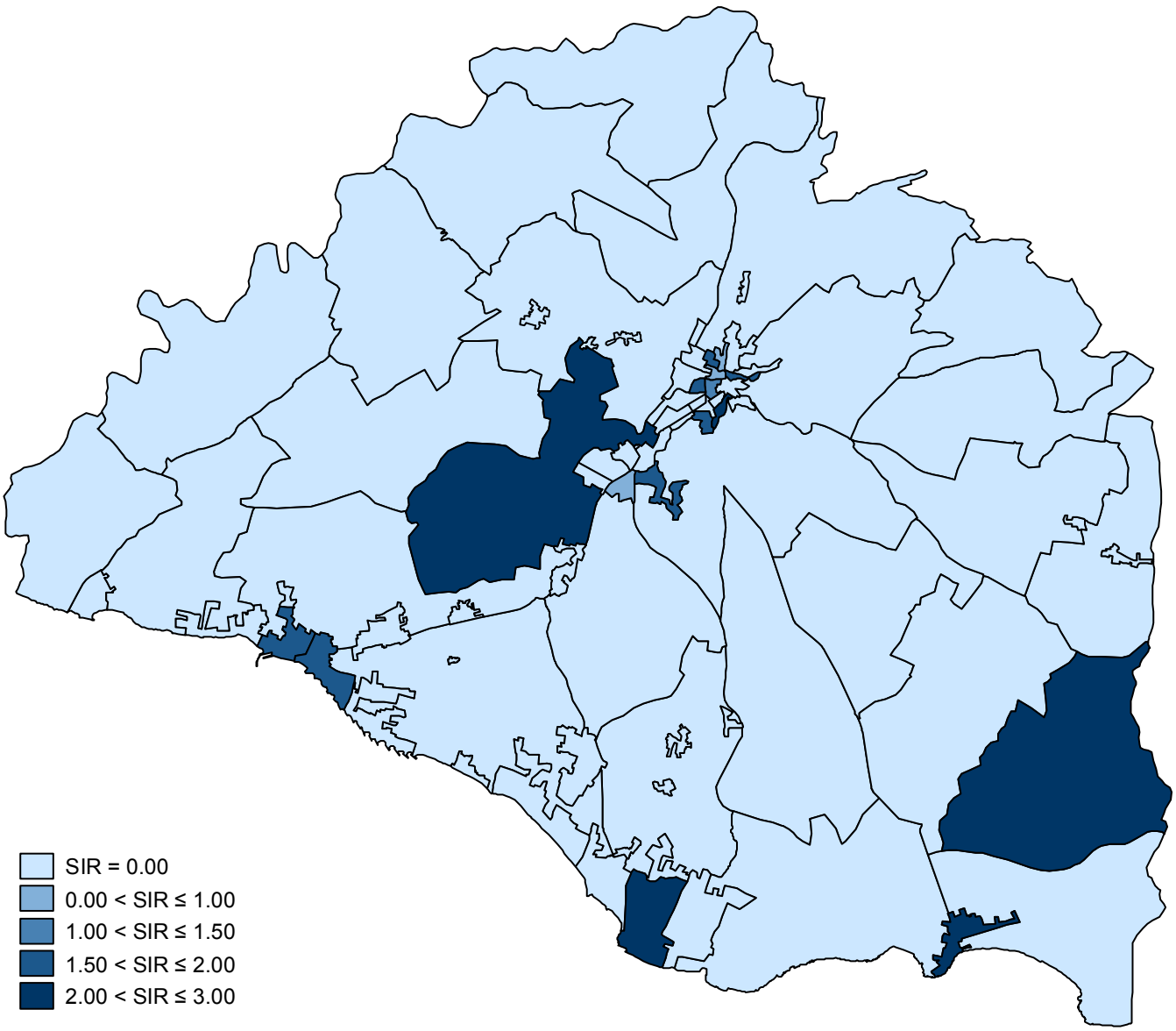
Zooming on Comiso municipality:



Zooming on Comiso municipality centre:



Zooming on Scicli municipality:



References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

ANNEX 16 – REPORT WITH ANALYSIS OF SIRACUSA CANCER REGISTRY DATA

From Siracusa CR data to analysis: first results

In October, the Siracusa CR provided the complete set of data, geo-coded at Census Tract (SEZ2001) level, totally covering the Siracusa province.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step-by-step discussion about the best model and software to be used for WASABY's spatial analysis, at Census Tract (named SEZ2001 for Siracusa province) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2004 to the 31st December 2013, for a total number of 587 female breast cancer patients aged 0-49 year. For 5 cases the attribution of the CT was not possible, therefore they were excluded from the analysis and 582 were considered.

The Italian National Statistics Office provided the 2001 shapefile at CT level for the Siracusa area, the 2001 Census Population and the INT provided the 2001 European Deprivation Index (EDI) at SEZ2001 level.

Merging observed 0-49 cases by Census tract with Siracusa province shapefile and 2001 EDI dataset

A dataset with SEZ2001 as key variable was built, merging the information from the Siracusa shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2001 dataset with the geo-coded Siracusa CR dataset.

The result was a dataset in which the rows reported the SEZ2001 with the sum of the observed cases (from the Siracusa CR) and all the other information at SEZ2001 level (from the other datasets).

The 582 considered incident cases were distributed in 358 SEZ2001s. In 1315 SEZ2001s where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Italian (`sir049f_ita`), European (`sir049f_eu`) and World (`sir049f_wor`) population standard were used.

The following table reports the average SIRs for the Siracusa province, along with the standard error and the 95% C.I.

```
Mean estimation                   Number of obs    =      1,673

------------------------------------------------------------------------------
             |       Mean   Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sir049f_ita |   .9255733   .1934298       .5461832    1.304963
  sir049f_eu |   1.079391   .2255751       .6369512     1.52183
 sir049f_wor |   1.581079   .3304199       .9329989    2.229159
------------------------------------------------------------------------------
```

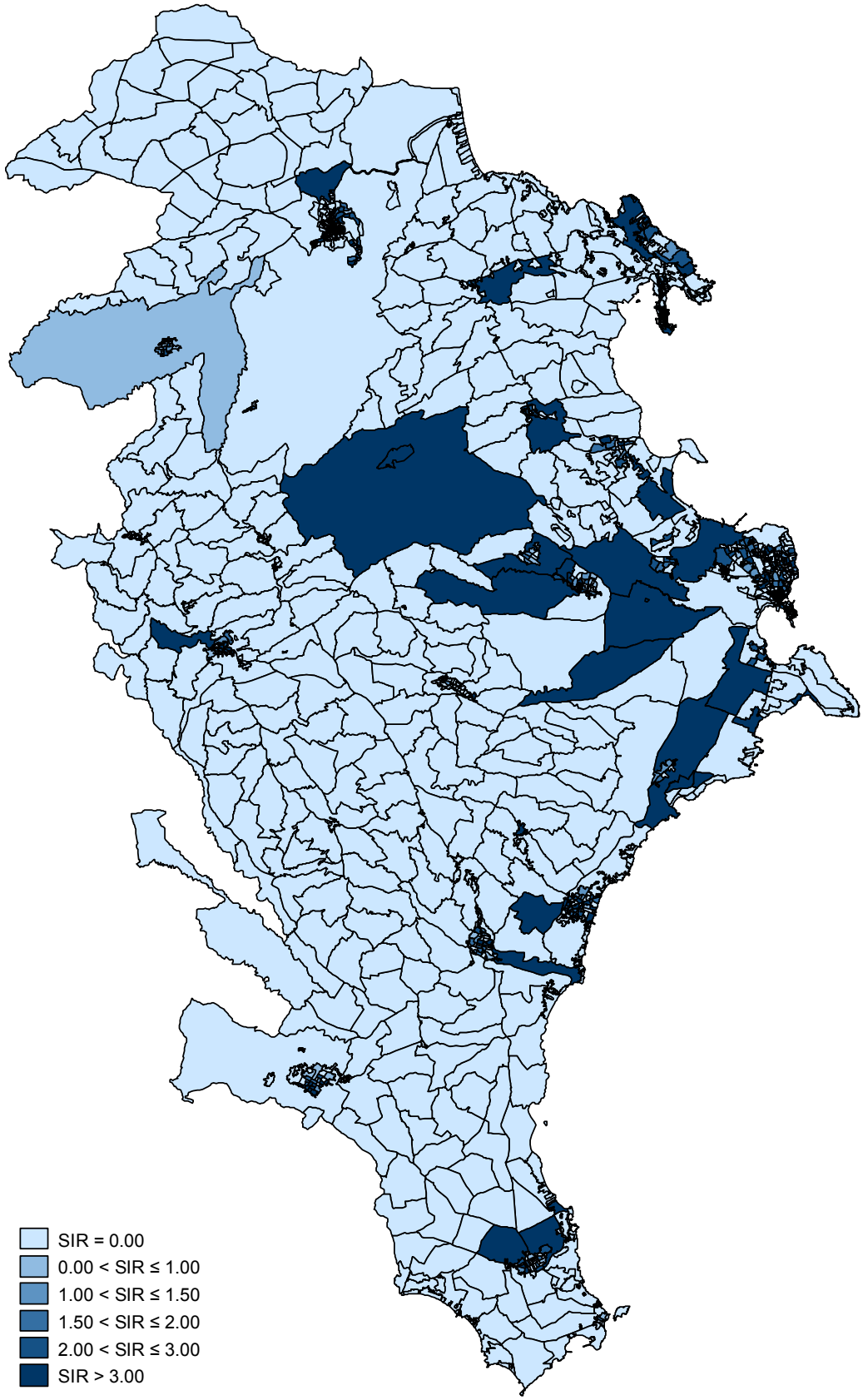
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
$0.00 < SIR \leq 1.00$
$1.00 < SIR \leq 1.50$
$1.50 < SIR \leq 2.00$
$2.00 < SIR \leq 3.00$
$SIR > 3.00$

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by World standardization was reported. No smoothing was applied.

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object sirmat
------------------------------------------
        Matrix |               Description
---------------+--------------------------
    Dimensions |               1673 x 1673
     Stored as |               1673 x 1673
 Links         |
         total |                     9626
           min |                        0
          mean |                 5.753736
           max |                       73
------------------------------------------
warning: spatial-weighting matrix contains 5 islands
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by DE 2011). The EDI 2001 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Italy

```
Spatial autoregressive model                    Number of obs   =      1673
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_ita |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_ita     |
zediweighted_01 |  -.0310411   .0649448    -0.48   0.633    -.1583306    .0962484
          _cons |   .5174883   .3211914     1.61   0.107    -.1120353    1.147012
----------------+-------------------------------------------------------------
lambda          |
          _cons |   4.773533   2.619499     1.82   0.068    -.3605908    9.907657
----------------+-------------------------------------------------------------
rho             |
          _cons |  -8.689676   3.480004    -2.50   0.013    -15.51036   -1.868995
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =      1673
(GS2SLS estimates)

------------------------------------------------------------------------------
     sir049f_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_eu      |
zediweighted_01 |  -.0361997   .0757377    -0.48   0.633    -.1846429    .1122436
          _cons |   .6034876   .3745689     1.61   0.107     -.130654    1.337629
----------------+-------------------------------------------------------------
lambda          |
          _cons |   4.773533   2.619499     1.82   0.068     -.360591    9.907658
----------------+-------------------------------------------------------------
rho             |
          _cons |  -8.689676   3.480004    -2.50   0.013    -15.51036   -1.868994
------------------------------------------------------------------------------
```

Standardization: World

Spatial autoregressive model — Number of obs = 1673
(GS2SLS estimates)

sir049f_wor	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sir049f_wor						
zediweighted_01	-.0530248	.1109398	-0.48	0.633	-.2704628	.1644131
_cons	.8839818	.5486643	1.61	0.107	-.1913805	1.959344
lambda						
_cons	4.773533	2.619499	1.82	0.068	-.3605911	9.907658
rho						
_cons	-8.689676	3.480004	-2.50	0.013	-15.51036	-1.868994

The graphical representation was reported only for the World standardization.

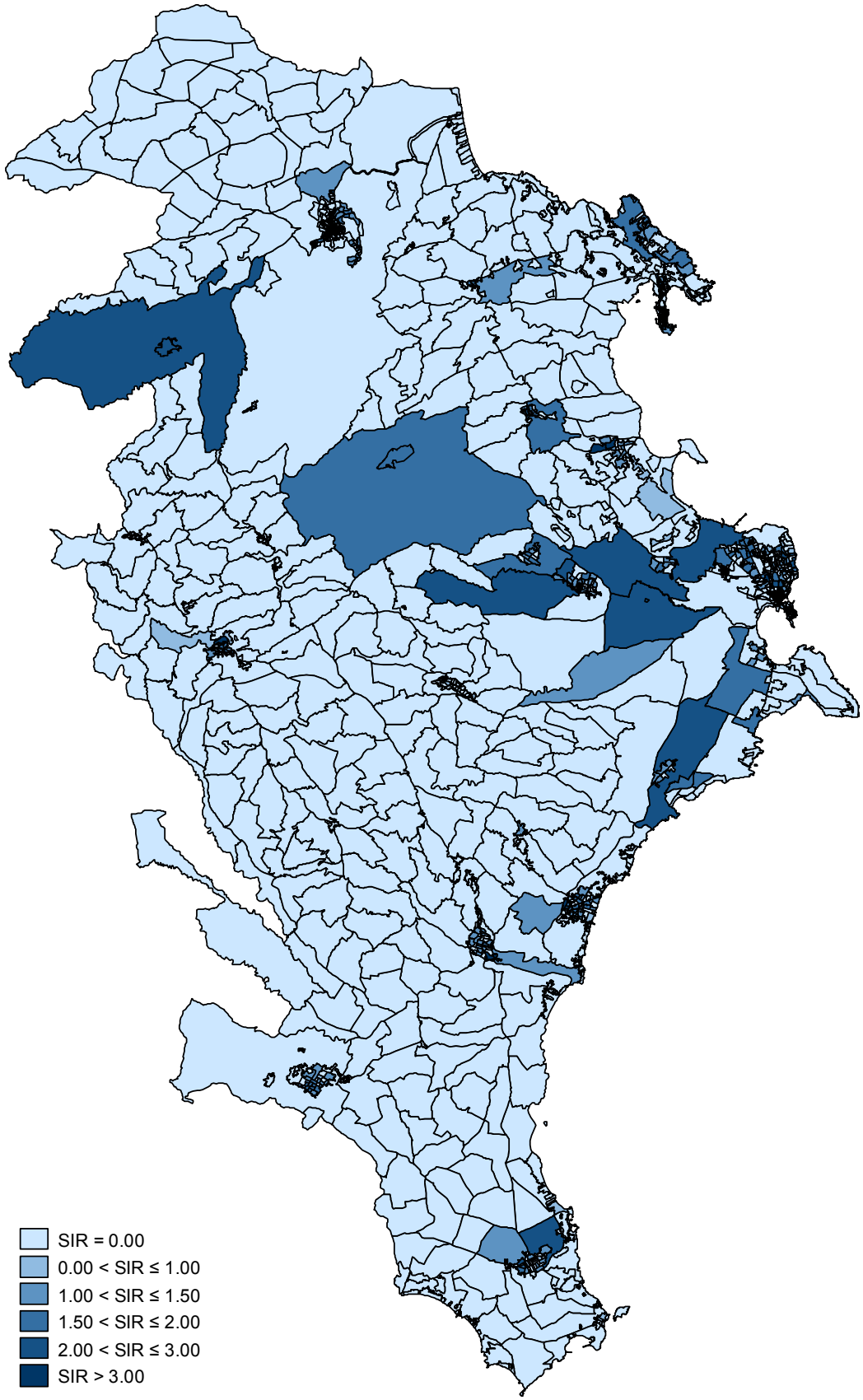
Moreover, also a zoom on the seven largest municipalities (Siracusa, Augusta, Avola, Lentini, Noto, Floridia, Pachino) was reported.

References

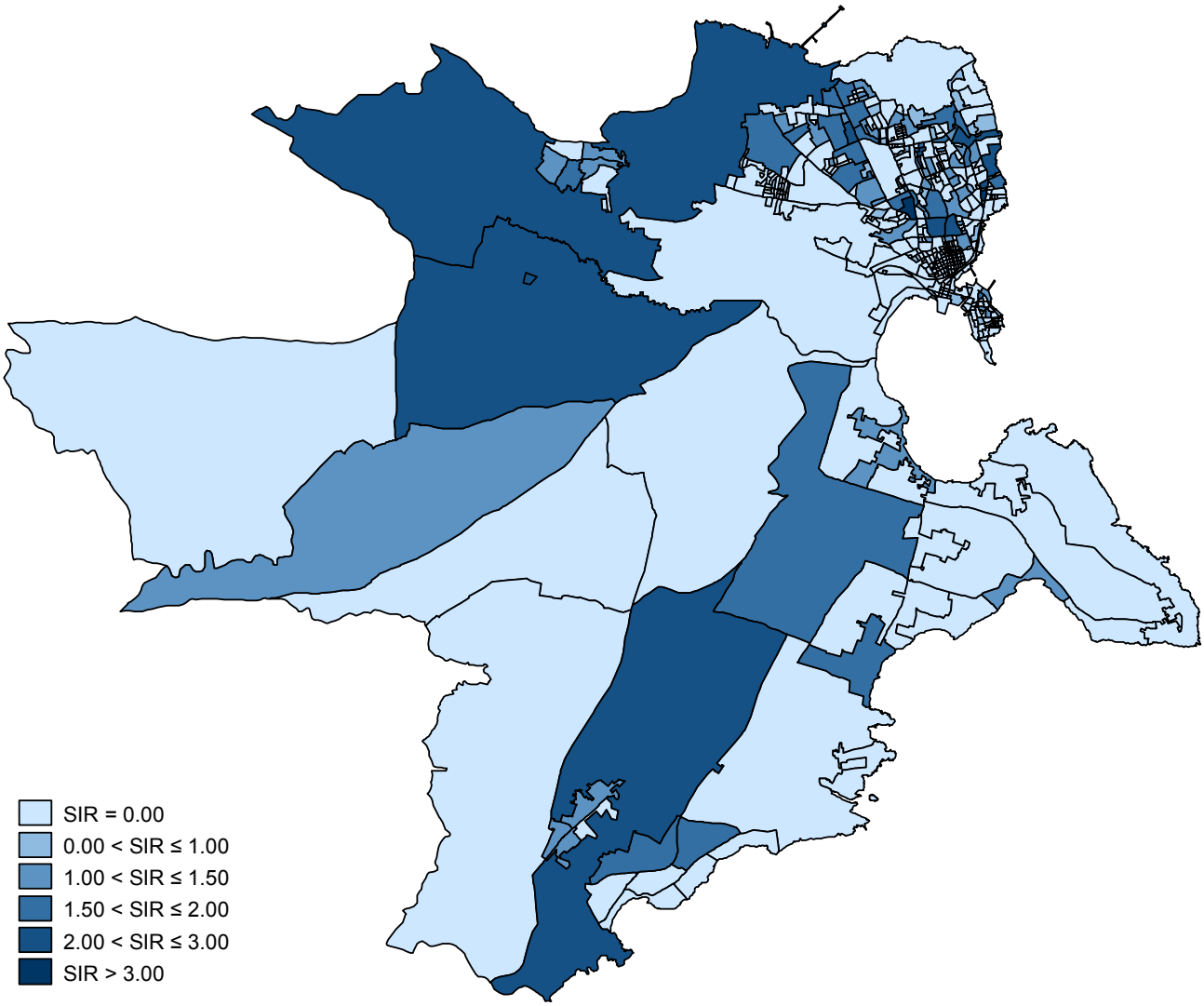
Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

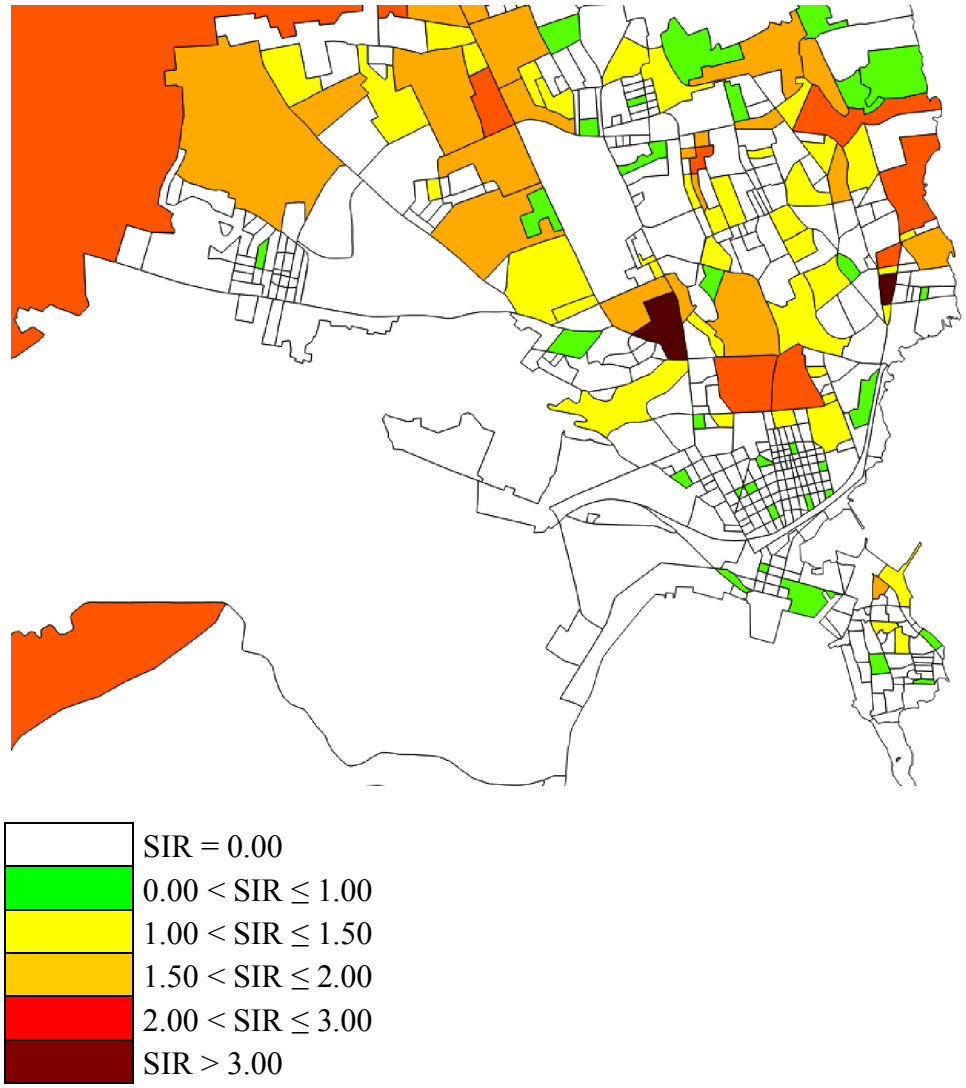
Standardization: World



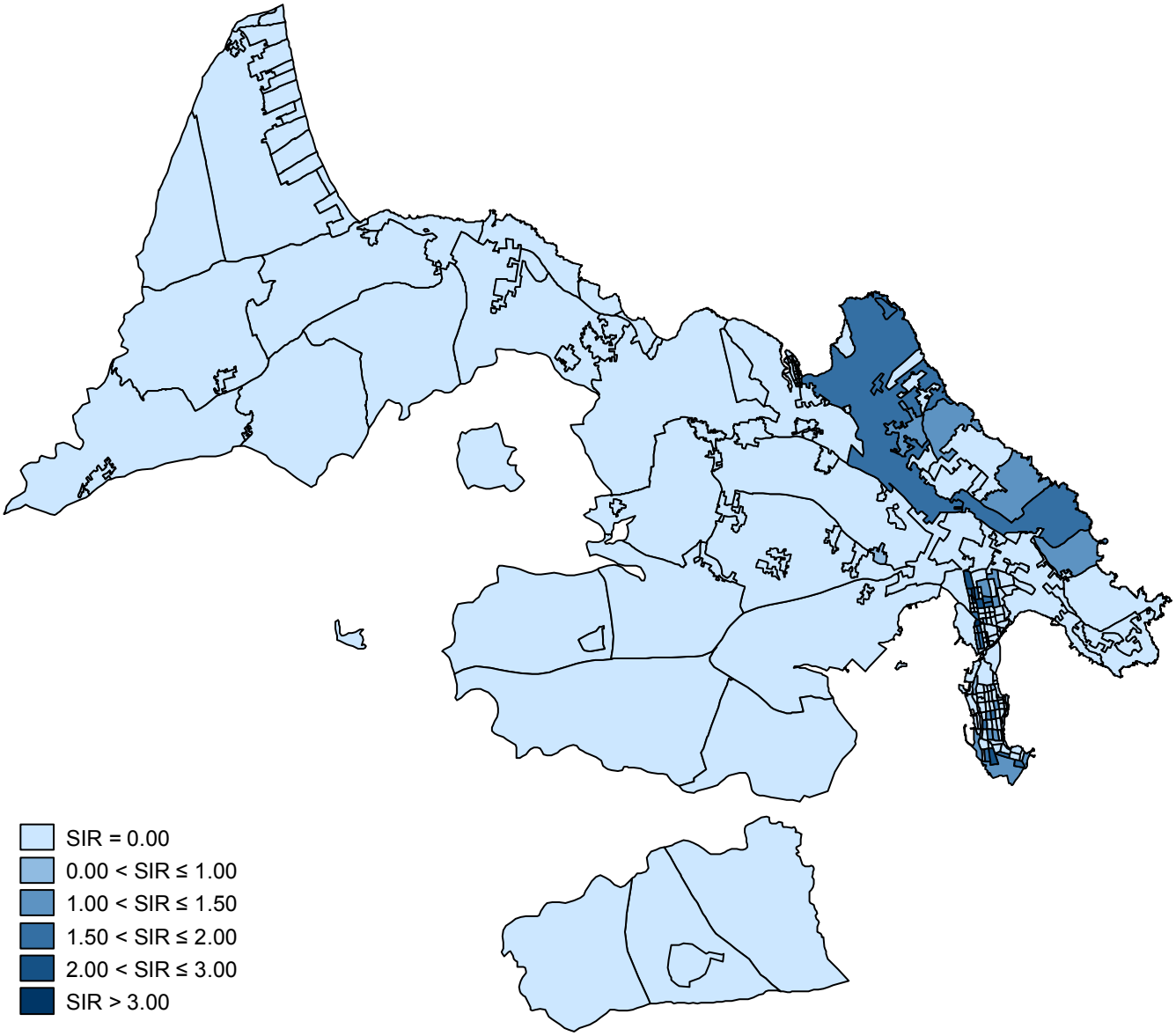
Zooming on Siracusa municipality:



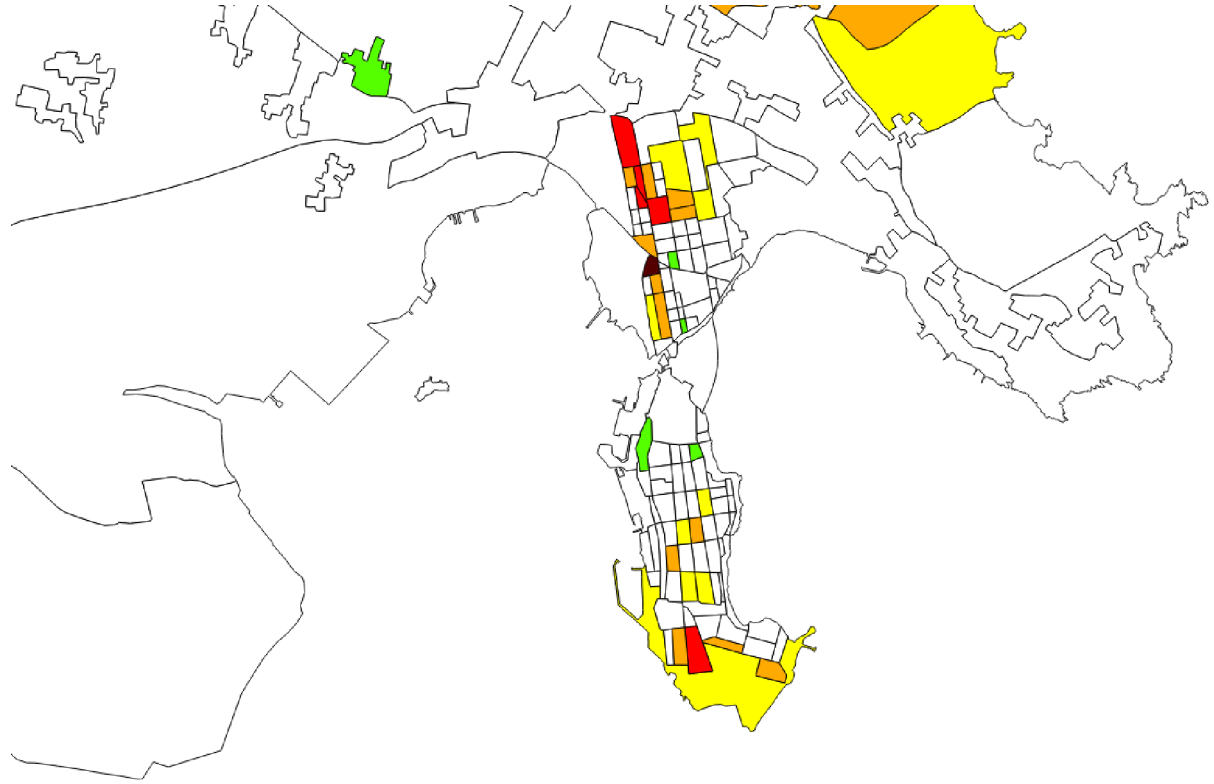
Focus on Siracusa city centre:



Zooming on Augusta municipality:

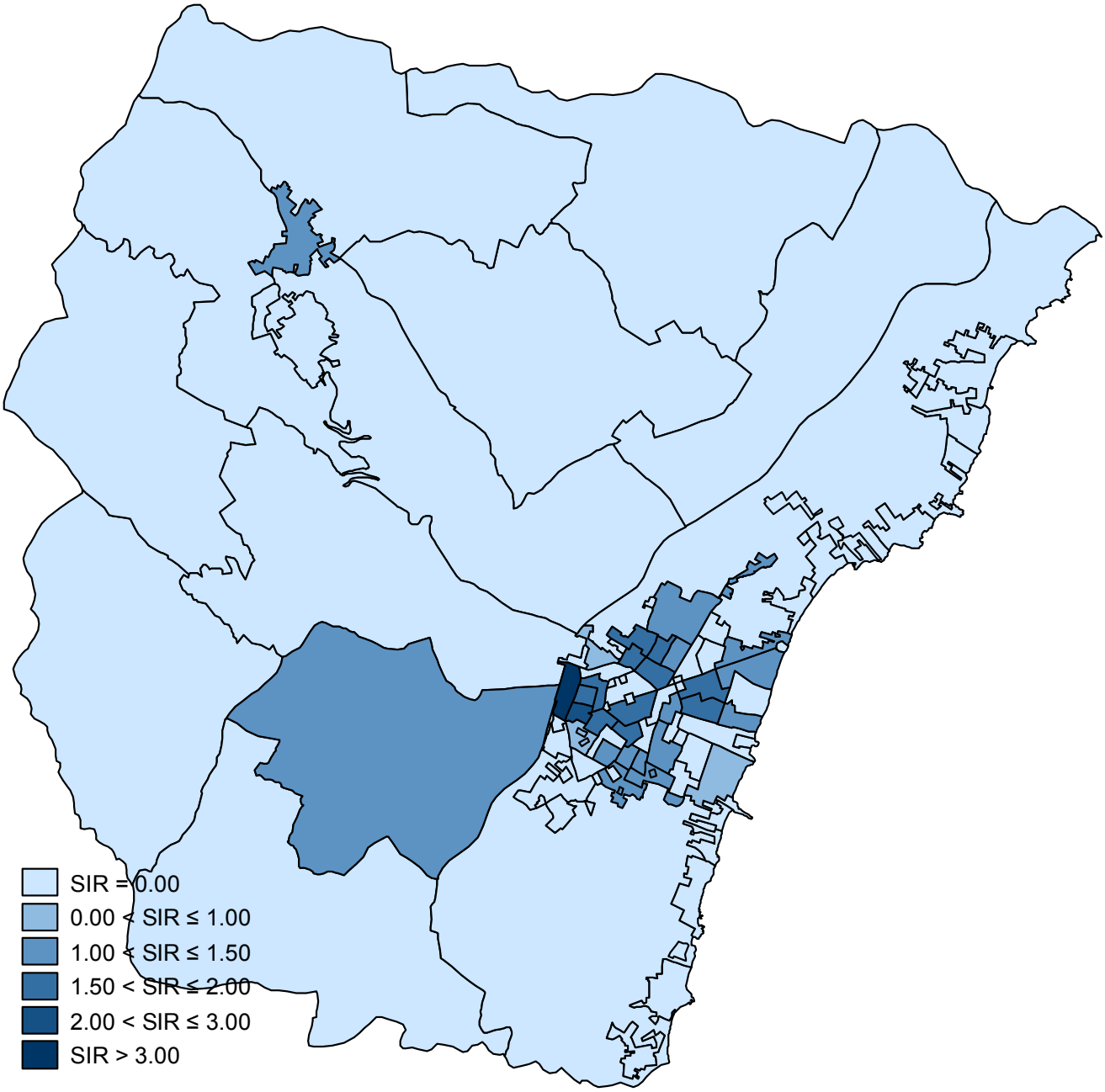


Focus on Augusta city centre:

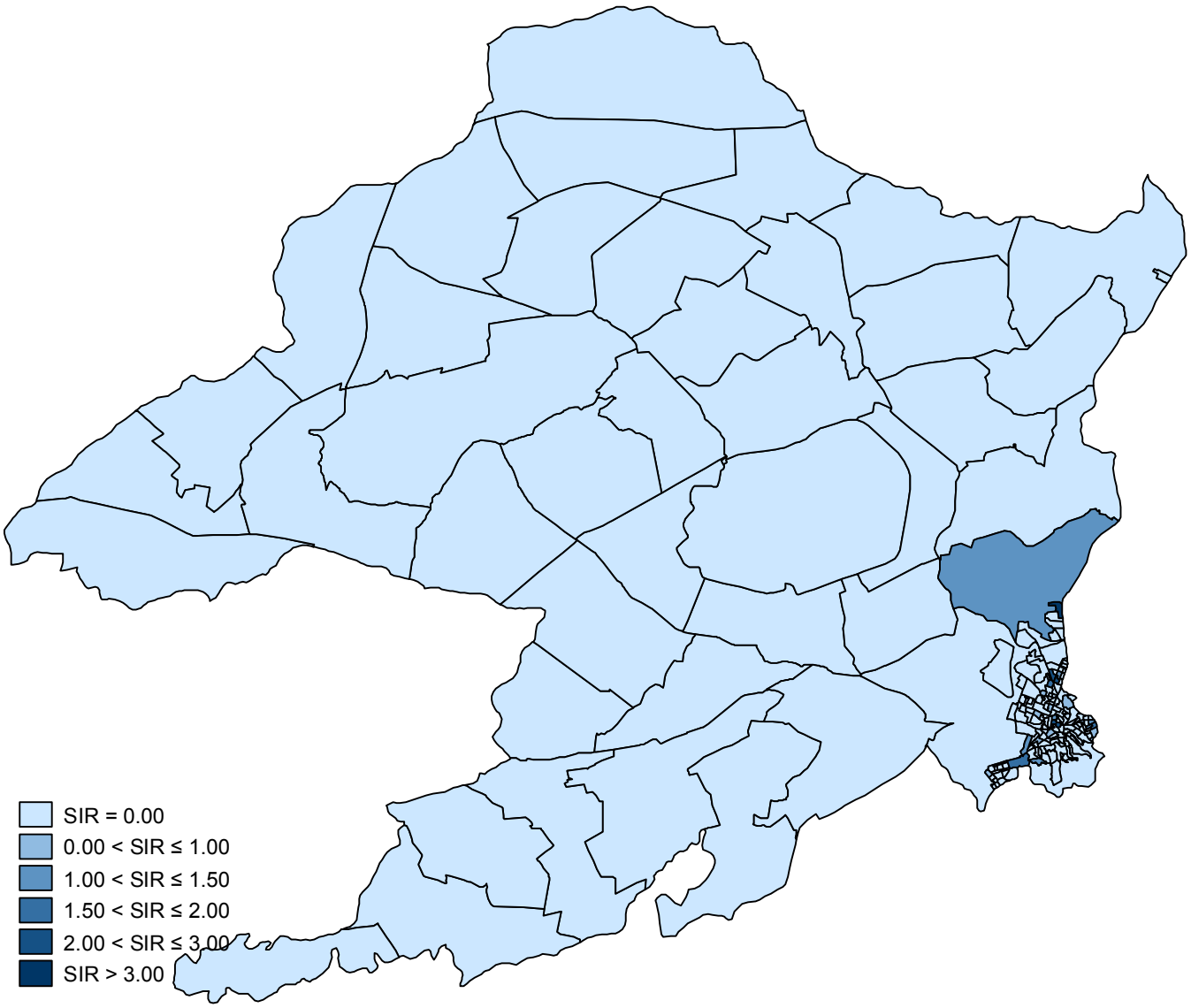


SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

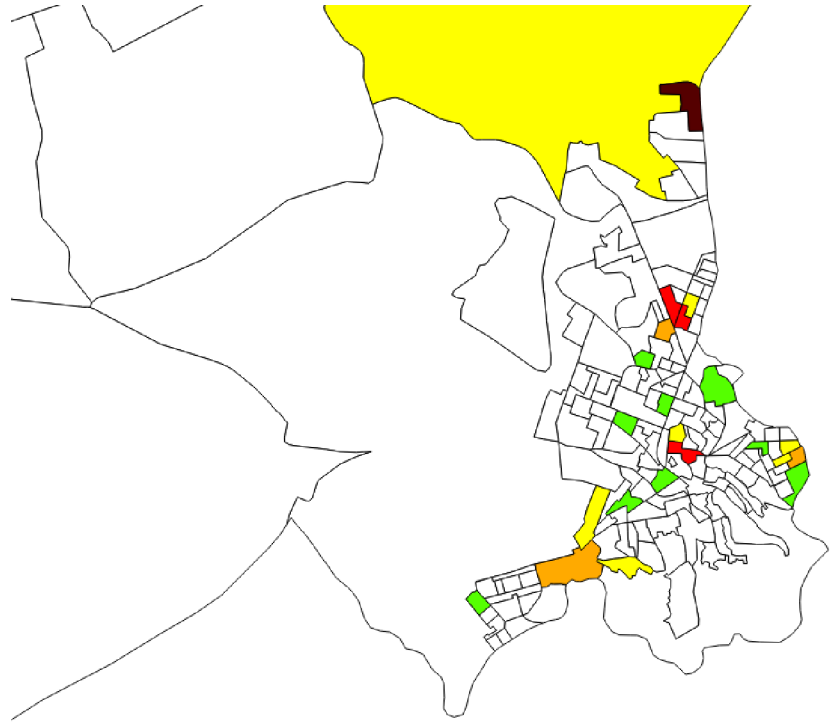
Zooming on Avola municipality:

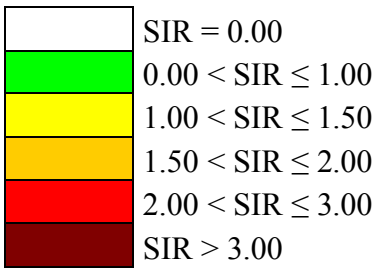


Zooming on Lentini municipality:

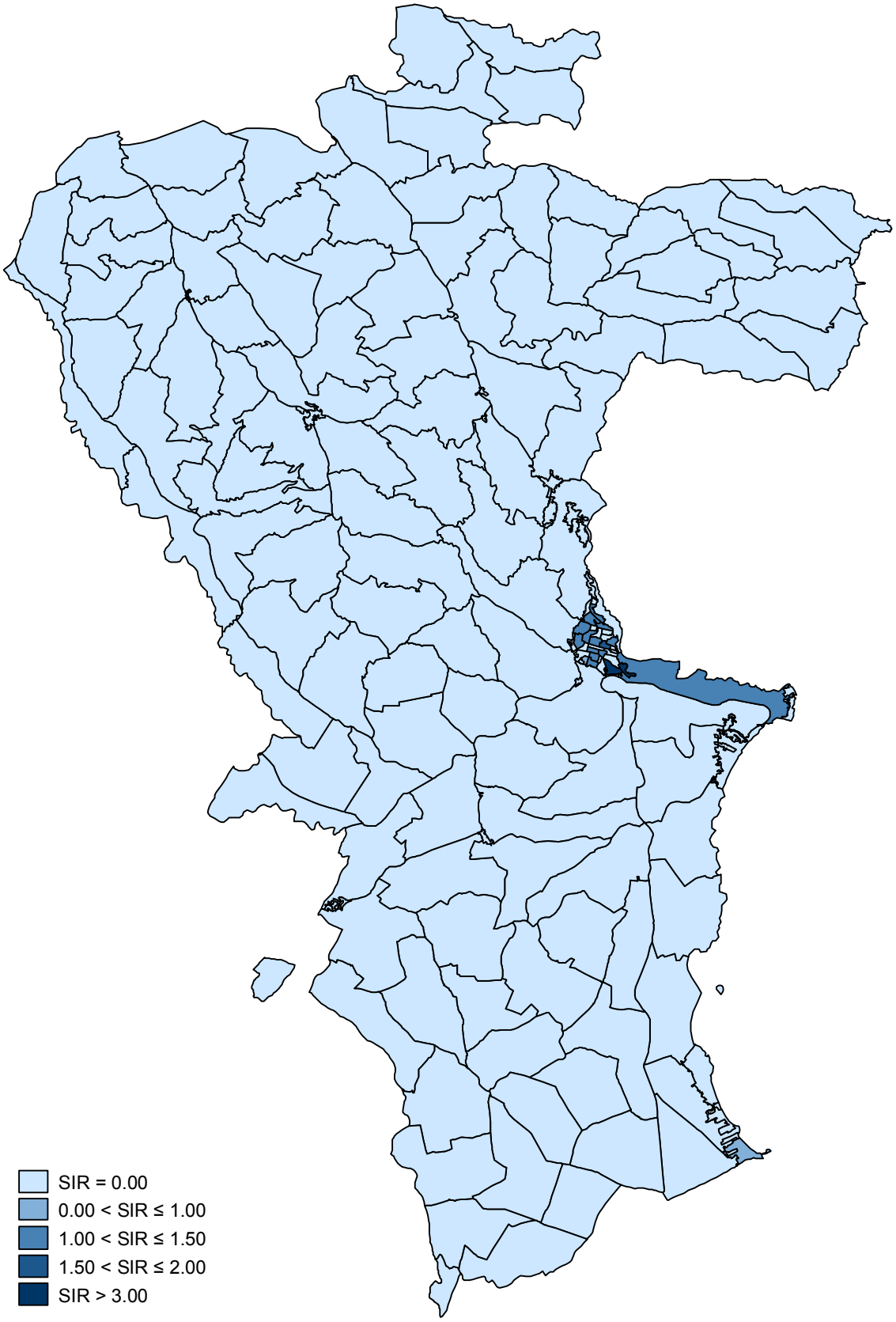


Focus on Lentini city centre:



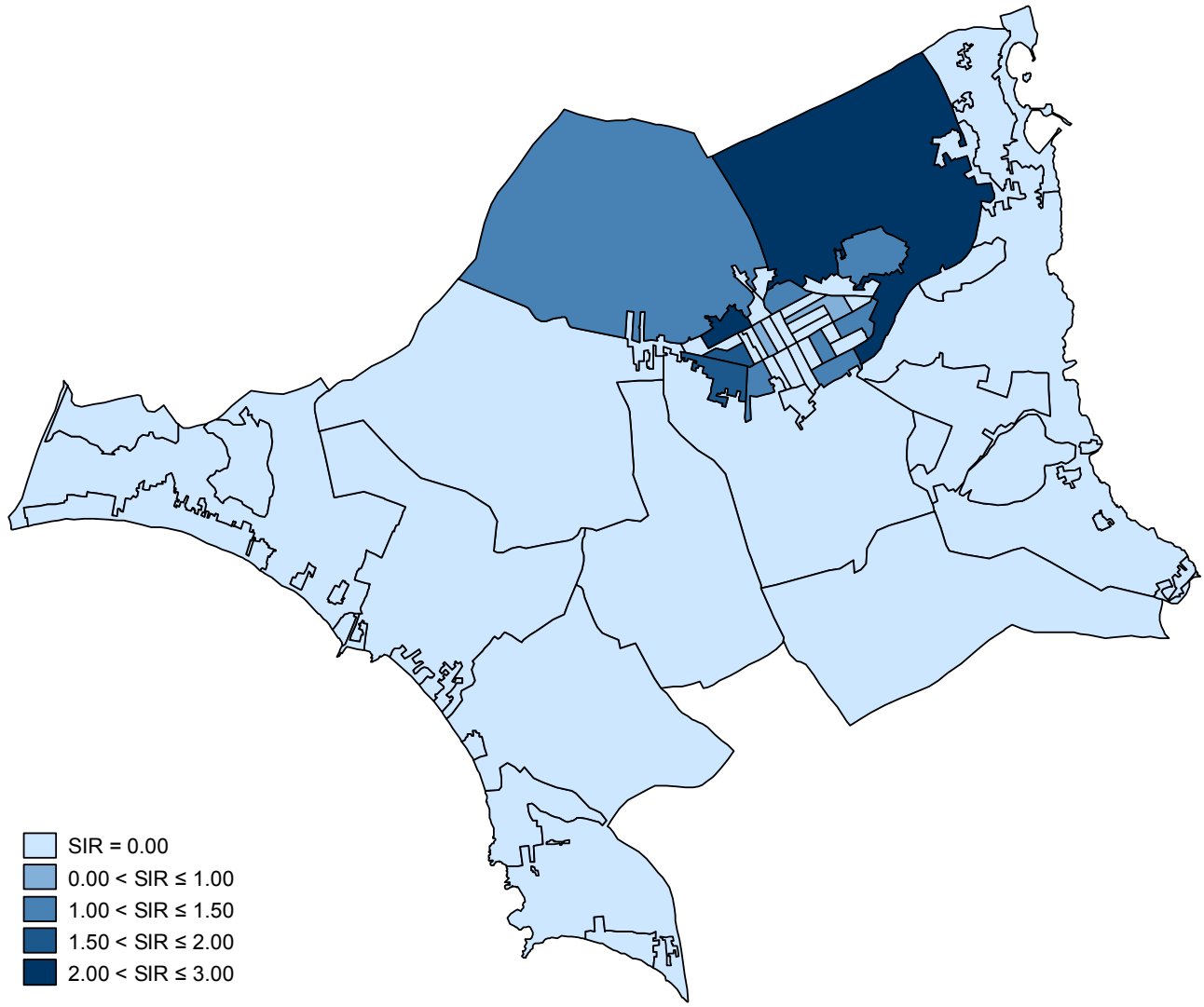


Zooming on Noto municipality:



Zooming on Floridia municipality:

Zooming on Pachino municipality:



ANNEX 17 – REPORT WITH ANALYSIS OF PARMA CANCER REGISTRY DATA

From Parma CR data to analysis: first results

In November, the Parma CR provided the complete set of data, geo-coded at Census Tract (SEZ2001) level, totally covering the Parma municipality.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step-by-step discussion about the best model and software to be used for WASABY's spatial analysis, at Census Tract (named SEZ2001 for Parma municipality) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2001 to the 31st December 2010, for a total number of 381 female breast cancer patients aged 0-49 year. For 7 cases the attribution of the CT was not possible, therefore they were excluded from the analysis and 374 were considered.

The Italian National Statistics Office provided the 2001 shapefile at CT level for the Parma municipality area, the 2001 Census Population and the INT provided the 2001 European Deprivation Index (EDI) at SEZ2001 level.

Merging observed 0-49 cases by Census tract with Parma municipality shapefile and 2001 EDI dataset

A dataset with SEZ2001 as key variable was built, merging the information from the Parma shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2001 dataset with the geo-coded Parma CR dataset.

The result was a dataset in which the rows reported the SEZ2001 with the sum of the observed cases (from the Parma CR) and all the other information at SEZ2001 level (from the other datasets).

The 374 considered incident cases were distributed in 274 SEZ2001s. In 972 SEZ2001s where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Italian (`sir049f_ita`), European (`sir049f_eu`) and World (`sir049f_wor`) population standard were used.

The following table reports the average SIRs for the Parma municipality, along with the standard error and the 95% C.I.

```
Mean estimation                   Number of obs    =      1,246

------------------------------------------------------------------------------
             |       Mean   Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sir049f_ita |   .7051901   .0569781       .5934064    .8169739
  sir049f_eu |   .8176226   .0660625       .6880166    .9472287
 sir049f_wor |   1.197165   .0967288       1.007395    1.386934
------------------------------------------------------------------------------
```

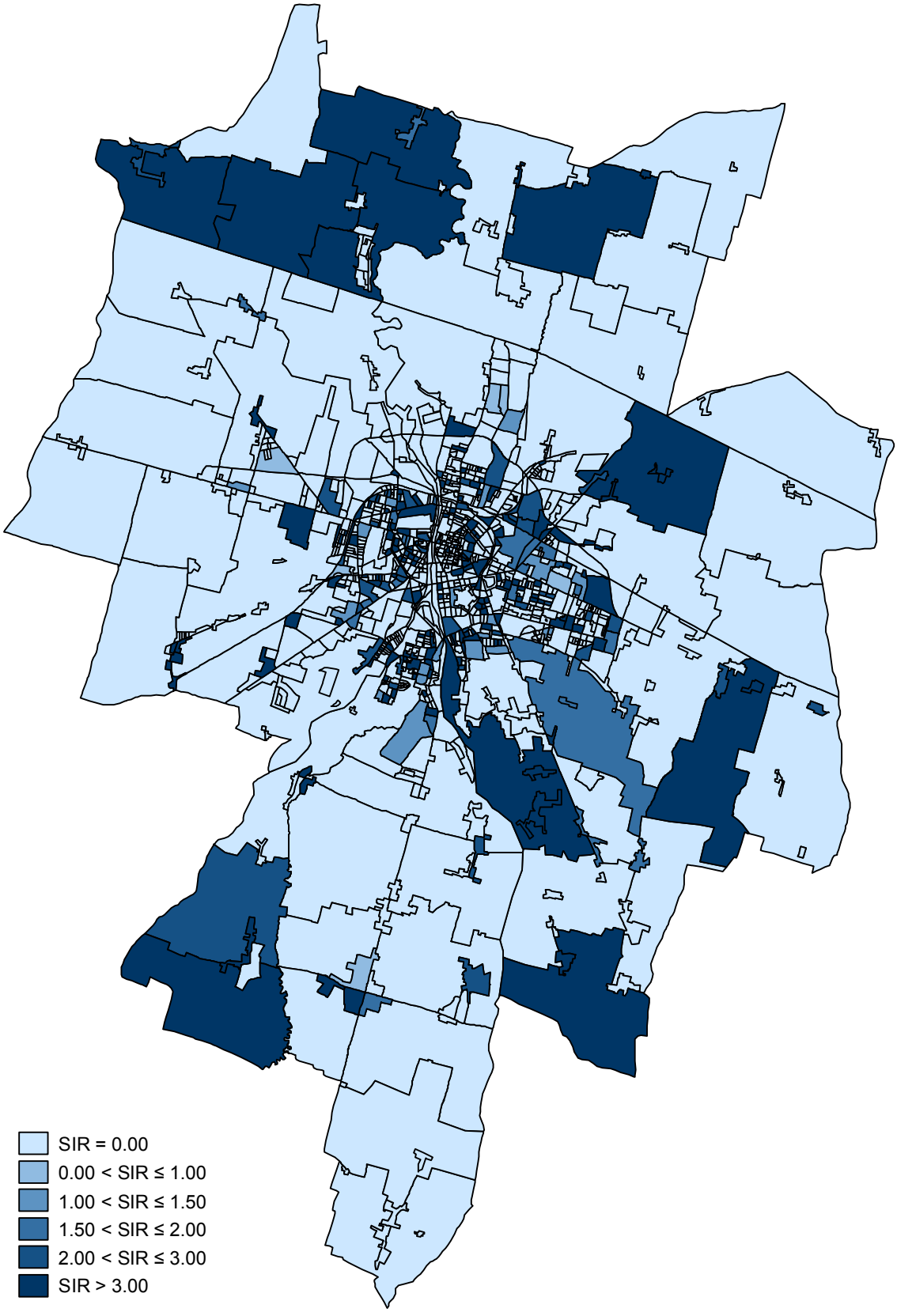
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by World standardization was reported. No smoothing was applied.

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object parmat
------------------------------------------
        Matrix |               Description
---------------+--------------------------
    Dimensions |               1246 x 1246
     Stored as |               1246 x 1246
 Links         |
         total |                     7858
           min |                        1
          mean |                 6.306581
           max |                       30
------------------------------------------
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by DE 2011). The EDI 2001 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Italy

```
Spatial autoregressive model                    Number of obs   =      1246
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_ita |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_ita     |
zediweighted_01 |   .0615893   .0328752     1.87   0.061    -.0028449    .1260236
          _cons |   .8708309   .4792856     1.82   0.069    -.0685517    1.810213
----------------+-------------------------------------------------------------
lambda          |
          _cons |  -.9918726   3.103977    -0.32   0.749    -7.075556    5.091811
----------------+-------------------------------------------------------------
rho             |
          _cons |  -1.038265    3.00141    -0.35   0.729     -6.92092    4.844389
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =      1246
(GS2SLS estimates)

------------------------------------------------------------------------------
     sir049f_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_eu      |
zediweighted_01 |   .0714088   .0381167     1.87   0.061    -.0032985    .1461162
          _cons |   1.009672   .5557009     1.82   0.069    -.0794812    2.098826
----------------+-------------------------------------------------------------
lambda          |
          _cons |  -.9918729   3.103977    -0.32   0.749    -7.075556    5.091811
----------------+-------------------------------------------------------------
rho             |
          _cons |  -1.038265   3.001409    -0.35   0.729    -6.920919    4.844389
------------------------------------------------------------------------------
```

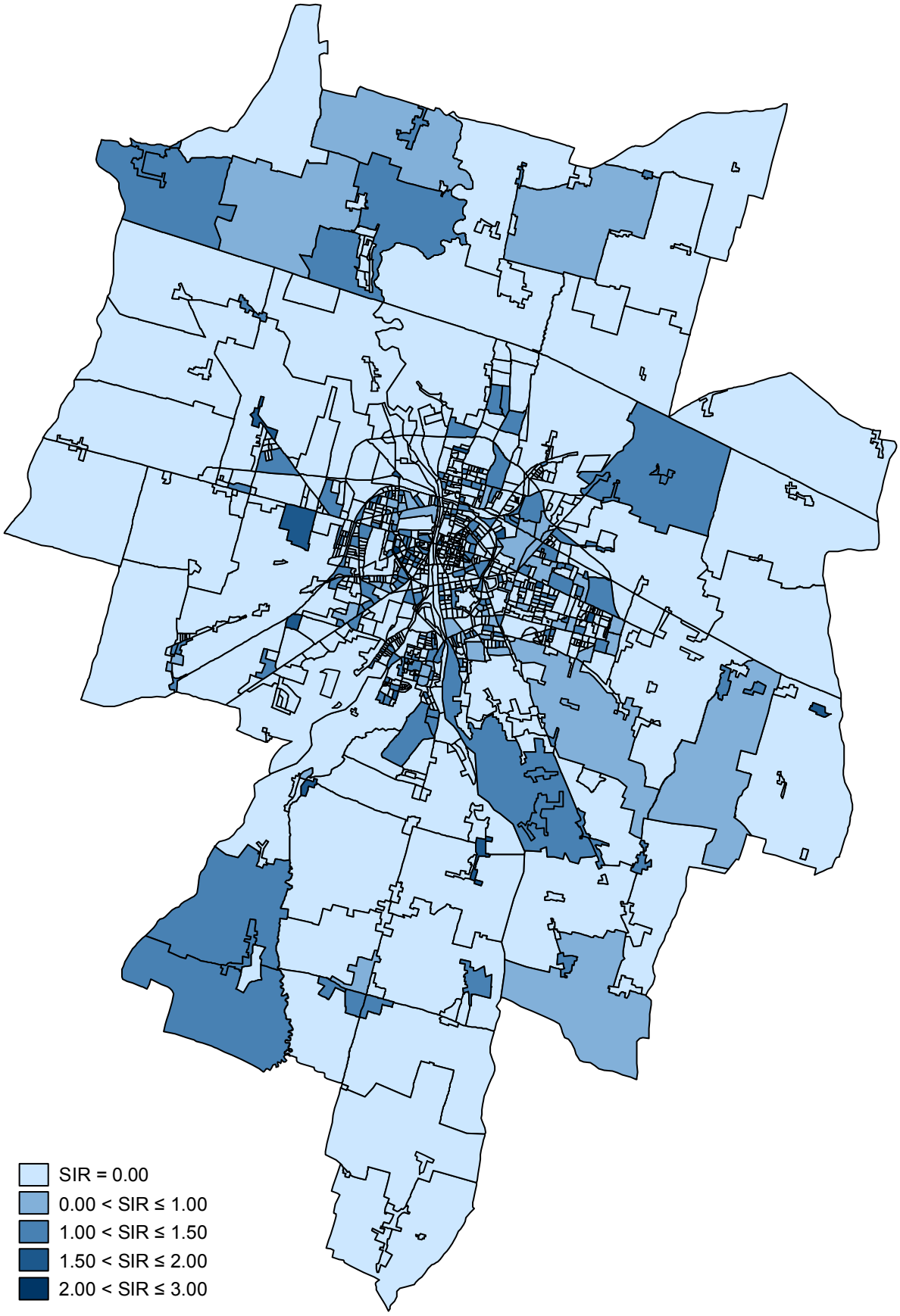
Standardization: World

```
Spatial autoregressive model                    Number of obs   =      1246
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_wor     |
zediweighted_01 |    .104557   .0558106     1.87   0.061    -.0048297    .2139436
          _cons |   1.478364   .8136583     1.82   0.069    -.1163765    3.073105
----------------+-------------------------------------------------------------
lambda          |
          _cons |  -.9918731   3.103977    -0.32   0.749    -7.075557    5.091811
----------------+-------------------------------------------------------------
rho             |
          _cons |  -1.038265   3.001409    -0.35   0.729    -6.920919    4.844389
------------------------------------------------------------------------------
```

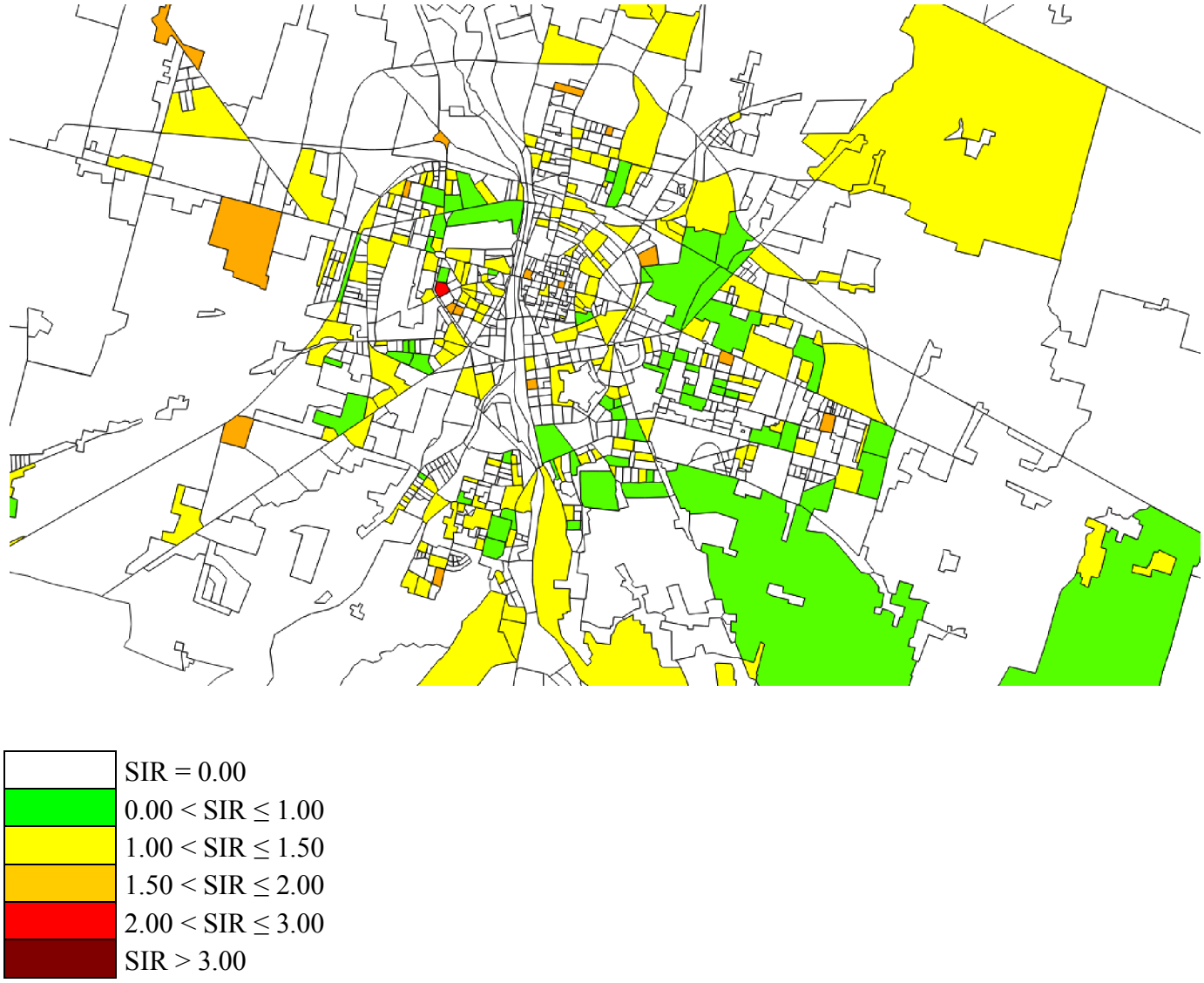
The graphical representation was reported only for the World standardization.

Moreover, also a zoom on the city centre was reported.

Standardization: World



Zooming on Parma city centre:



References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

ANNEX 18 – REPORT WITH ANALYSIS OF NAPOLI 3 SOUTH CANCER REGISTRY DATA

From Napoli CR data to analysis: first results

In January, the Napoli CR provided the complete set of data, geo-coded at Census Tract (SEZ2001 and SEZ2011) level, totally covering the ASL3 Napoli South area.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step-by-step discussion about the best model and software to be used for WASABY's spatial analysis, at Census Tract (named SEZ2001 for ASL3 Napoli South area) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2008 to the 31st December 2015, for a total number of 1514 female breast cancer patients aged 0-49 year. For 17 (1.1%) cases the attribution of the CT was not possible, therefore they were excluded from the analysis and 1497 were considered.

The Italian National Statistics Office provided the 2001 and 2011 shapefiles at CT level for the ASL3 Napoli South area, the 2001 and 2011 Census Population and the INT provided the 2001 European Deprivation Index (EDI) at SEZ2001 level. At the moment, analysis was developed for the 2001 geo-coding.

Merging observed 0-49 cases by Census tract with ASL3 Napoli South area shapefile and 2001 EDI dataset

A dataset with SEZ2001 as key variable was built, merging the information from the ASL3 Napoli South area shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2001 dataset with the geo-coded (x&y coordinates) ASL3 Napoli South area CR dataset.

The result was a dataset in which the rows reported the SEZ2001 with the sum of the observed cases (from the ASL3 Napoli South area CR) and all the other information at SEZ2001 level (from the other datasets).

The 1497 considered incident cases were distributed in 886 SEZ2001s. In 2425 SEZ2001s where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Italian (`sir049f_ita`), European (`sir049f_eu`) and World (`sir049f_wor`) population standard were used.

The following table reports the average SIRs for the ASL3 Napoli South area province, along with the standard error and the 95% C.I.

```
Mean estimation                   Number of obs    =      3,311

------------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+----------------------------------------------------
 sir049f_ita |   1.259182   .1207135      1.022502    1.495863
  sir049f_eu |   1.267484   .1215094      1.029243    1.505725
 sir049f_wor |   1.267494   .1215103      1.029251    1.505737
------------------------------------------------------------------
```

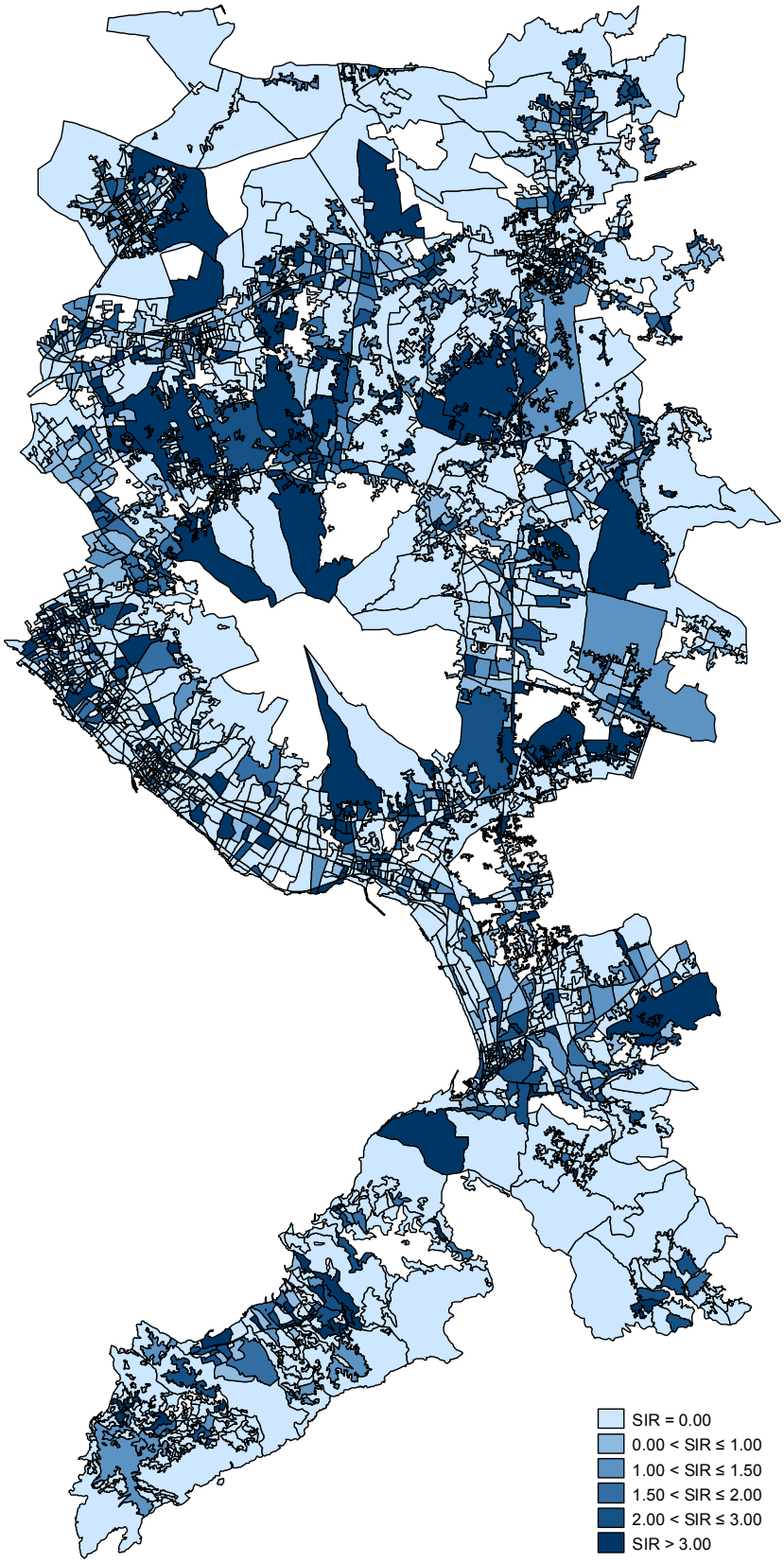
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by World standardization was reported. No smoothing was applied.

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object napmat
------------------------------------------
        Matrix |               Description
---------------+--------------------------
    Dimensions |               3311 x 3311
     Stored as |               3311 x 3311
 Links         |
         total |                    18370
           min |                        0
          mean |                 5.548173
           max |                       93
------------------------------------------
warning: spatial-weighting matrix contains 16 islands
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by EDI 2001). The EDI 2001 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Italy

```
Spatial autoregressive model                    Number of obs   =      3311
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_ita |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_ita     |
zediweighted_01 |  -.0056554   .0435858    -0.13   0.897     -.091082    .0797712
          _cons |   .5875425   .2414539     2.43   0.015     .1143015    1.060783
----------------+-------------------------------------------------------------
lambda          |
          _cons |   7.734503   2.399867     3.22   0.001     3.030849    12.43816
----------------+-------------------------------------------------------------
rho             |
          _cons |  -9.251664   3.156402    -2.93   0.003     -15.4381   -3.065229
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =      3311
(GS2SLS estimates)

------------------------------------------------------------------------------
     sir049f_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_eu      |
zediweighted_01 |  -.0056927   .0438731    -0.13   0.897    -.0916825    .0802971
          _cons |   .5914161   .2430458     2.43   0.015     .1150551    1.067777
----------------+-------------------------------------------------------------
lambda          |
          _cons |   7.734503   2.399867     3.22   0.001     3.030849    12.43816
----------------+-------------------------------------------------------------
rho             |
          _cons |  -9.251664   3.156402    -2.93   0.003     -15.4381   -3.065229
------------------------------------------------------------------------------
```

Standardization: World

```
Spatial autoregressive model                      Number of obs   =      3311
(GS2SLS estimates)

------------------------------------------------------------------------------
    sir049f_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
sir049f_wor     |
zediweighted_01 |  -.0056927   .0438735    -0.13   0.897    -.0916832    .0802977
          _cons |   .5914207   .2430477     2.43   0.015      .115056    1.067785
----------------+-------------------------------------------------------------
lambda          |
          _cons |   7.734503   2.399867     3.22   0.001     3.030849    12.43816
----------------+-------------------------------------------------------------
rho             |
          _cons |  -9.251664   3.156402    -2.93   0.003     -15.4381   -3.065229
------------------------------------------------------------------------------
```

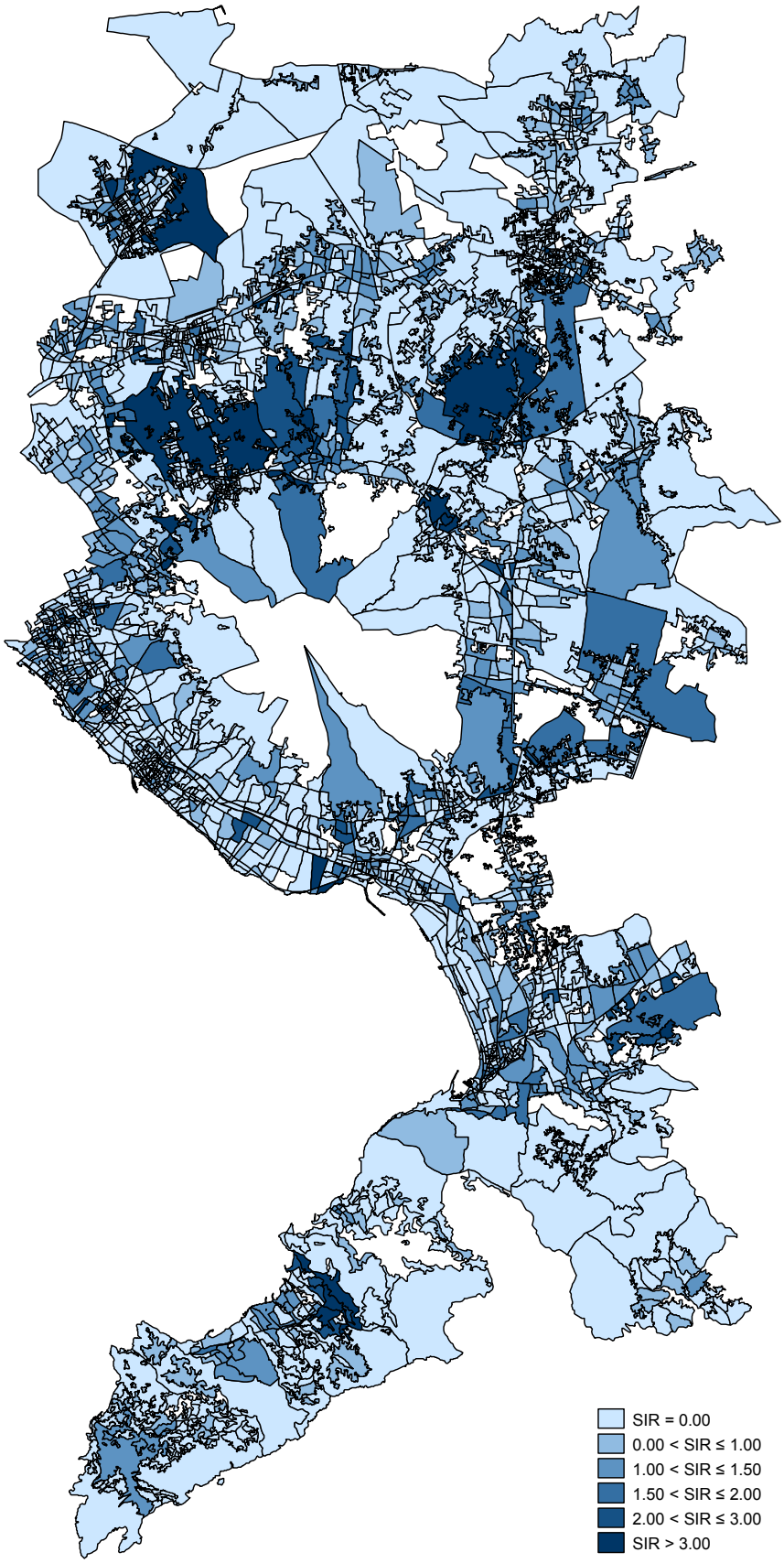
The graphical representation was reported only for the World standardization.

References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

Standardization: World



From Northern Portugal CR data to analysis: first results

On the end of November, the Northern Portugal (NP) CR provided the complete set of data, geo-coded at Parish (Freguesia) level.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step by step discussion on the best model and software to be used for WASABY's spatial analysis, at Parish (DICOFRE[3]) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2003 to the 31st December 2012, for a total number of 5561 female breast cancer patients aged 0-49 year.

The NP CR provided also the 2001 and 2011 Census Population, the WP5 provided the 2011 European Deprivation Index at DICOFRE level, while the 2011 shapefile at DICOFRE level for the NP was downloaded from the Portuguese National Geographic and Territorial Office website[4].

For 145 cases (2.6% of the total cases), the geo-coded information didn't match with the shapefile. They were excluded from the analysis.

Merging observed 0-49 cases by Parish with Northern Portugal province shapefile and EDI 2011 dataset

A dataset with DICOFRE as key variable was built, merging the information from the NP province shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2011 dataset with the geo-coded NP CR dataset.
The result was a dataset in which the rows reported the DICOFREs with the sum of the observed cases (from the NP CR) and all the other information at DICOFRE level (from the other datasets).

The 5416 considered incident cases were distributed in 1080 DICOFREs. In 961 DICOFREs where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

[3] DICOFRE is the variable label name used in the dataset to identify the Parishes variable.

[4] http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/carta_administrativa_oficial_de_portugal_caop_/caop__download_/carta_administrativa_oficial_de_portugal___versao_2011_2/

```
Result                              # of obs.
    ----------------------------------------
    not matched                           961
    matched                              1080
    total                                2041
    ----------------------------------------
```

After this action, all 2041 DICOFREs of the NP were allowed to be used in the SARAR models.

Computing the expected cases and the SIRs by DICOFREs

The expected cases and the SIRs were computed by direct standardization. World population standard was used (`sir049_wor`).

The following table reports the average SIRs for the NP, along with the standard error and the 95% C.I.

```
------------------------------------------------------------------
             |       Mean   Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------------
 sir049f_wor |   1.251441   .0502916      1.152813    1.350069
------------------------------------------------------------------
```

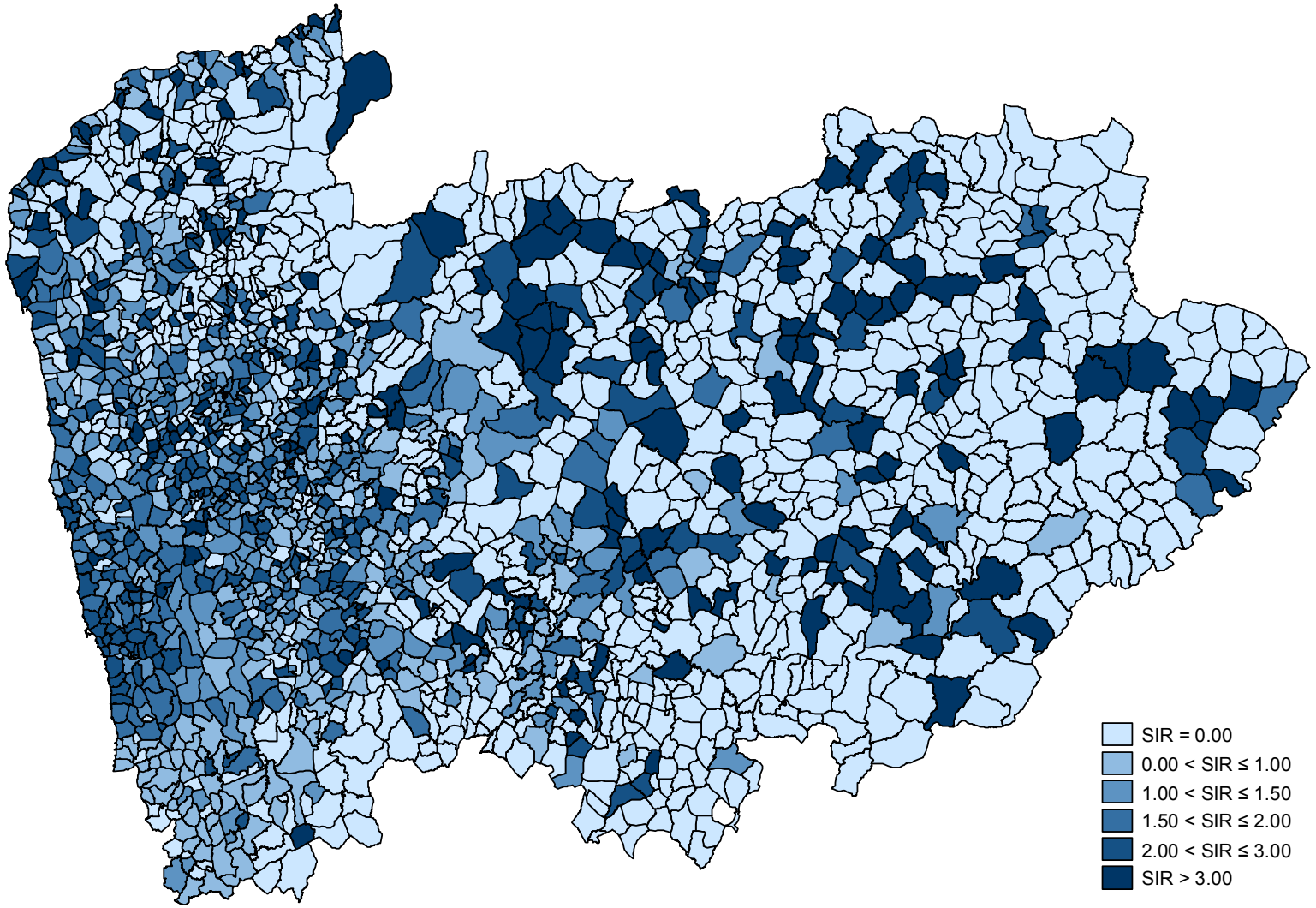
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

The map representing the SIRs by World standardization was reported. No smoothing was applied.

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object norpmat
-------------------------------------------
        Matrix |                 Description
---------------+---------------------------
    Dimensions |                 2041 x 2041
     Stored as |                 2041 x 2041
 Links         |
         total |                       11976
           min |                           0
          mean |                    5.867712
           max |                          15
-------------------------------------------
warning: spatial-weighting matrix contains 1 island
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by EDI 2011). The EDI 2011 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

Standardization: World

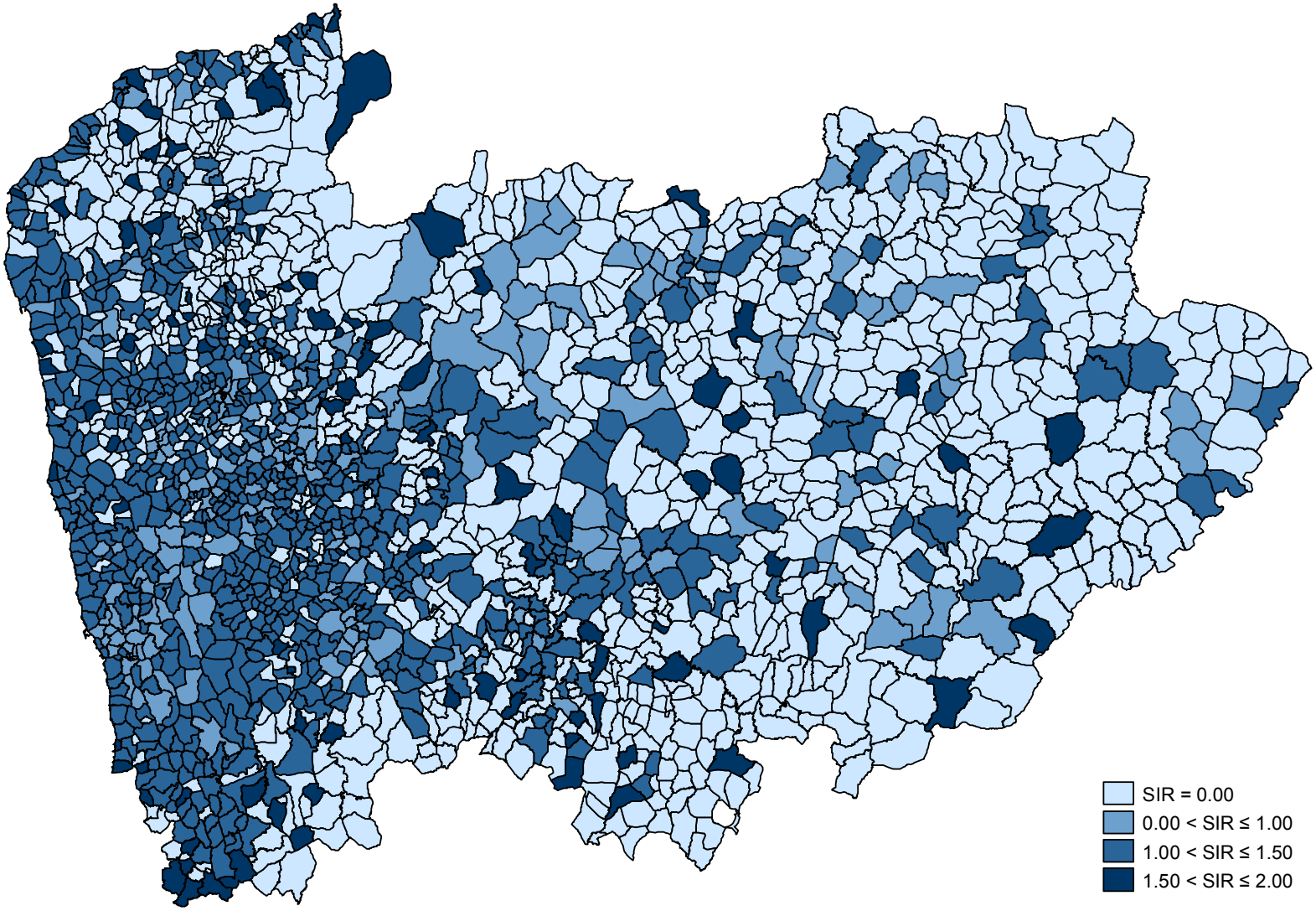
```
Spatial autoregressive model                    Number of obs   =      2041
(GS2SLS estimates)

------------------------------------------------------------------------------
 sir049f_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049f_wor  |
   EDI_SCORE |  -.0159531   .0303445    -0.53   0.599    -.0754272     .043521
       _cons |   1.622507   1.996062     0.81   0.416    -2.289703    5.534716
-------------+----------------------------------------------------------------
lambda       |
       _cons |  -.8239593   4.283189    -0.19   0.847    -9.218855    7.570936
-------------+----------------------------------------------------------------
rho          |
       _cons |   1.114992   .3116997     3.58   0.000     .5040713    1.725912
------------------------------------------------------------------------------
```

References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

Standardization: World



ANNEX 20 – REPORT WITH ANALYSIS OF CENTRAL PORTUGAL CANCER REGISTRY DATA

From Central Portugal CR data to analysis: first results

On the end of March 2019, the Central Portugal (CP) CR provided the complete set of data, geo-coded at Parish (Freguesia) level.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step by step discussion on the best model and software to be used for WASABY's spatial analysis, at Parish (DICOFRE[5]) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2001 to the 31st December 2011, for a total number of 3715 female breast cancer patients aged 0-49 year.

The NP CR provided also the 2001 and 2011 Census Population, the WP5 provided the 2011 European Deprivation Index at DICOFRE level, while the 2011 shapefile at DICOFRE level for the CP was downloaded from the Portuguese National Geographic and Territorial Office website[6].

All the cases were correctly geo-coded and used in the analyses.

Merging observed 0-49 cases by Parish with Central Portugal province shapefile and EDI 2011 dataset

A dataset with DICOFRE as key variable was built, merging the information from the CP province shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2011 dataset with the geo-coded CP CR dataset.
The result was a dataset in which the rows reported the DICOFREs with the sum of the observed cases (from the CP CR) and all the other information at DICOFRE level (from the other datasets).

The 3715 considered incident cases were distributed in 828 DICOFREs. In 627 DICOFREs where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

After this action, all 1455 DICOFREs of the CP were allowed to be used in the SARAR models.

[5] DICOFRE is the variable label name used in the dataset to identify the Parishes variable.

[6] http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/carta_administrativa_oficial_de_portugal_caop_/caop__download_/carta_administrativa_oficial_de_portugal___versao_2011_2/

Computing the expected cases and the SIRs by DICOFREs

The expected cases and the SIRs were computed by direct standardization. World population standard was used (`sir049_wor`).

The following table reports the average SIRs for the CP, along with the standard error and the 95% C.I.

```
Mean estimation                   Number of obs    =      1,455

------------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+----------------------------------------------------
 sir049f_wor |   .8542185   .0316243      .7921843    .9162527
------------------------------------------------------------------
```

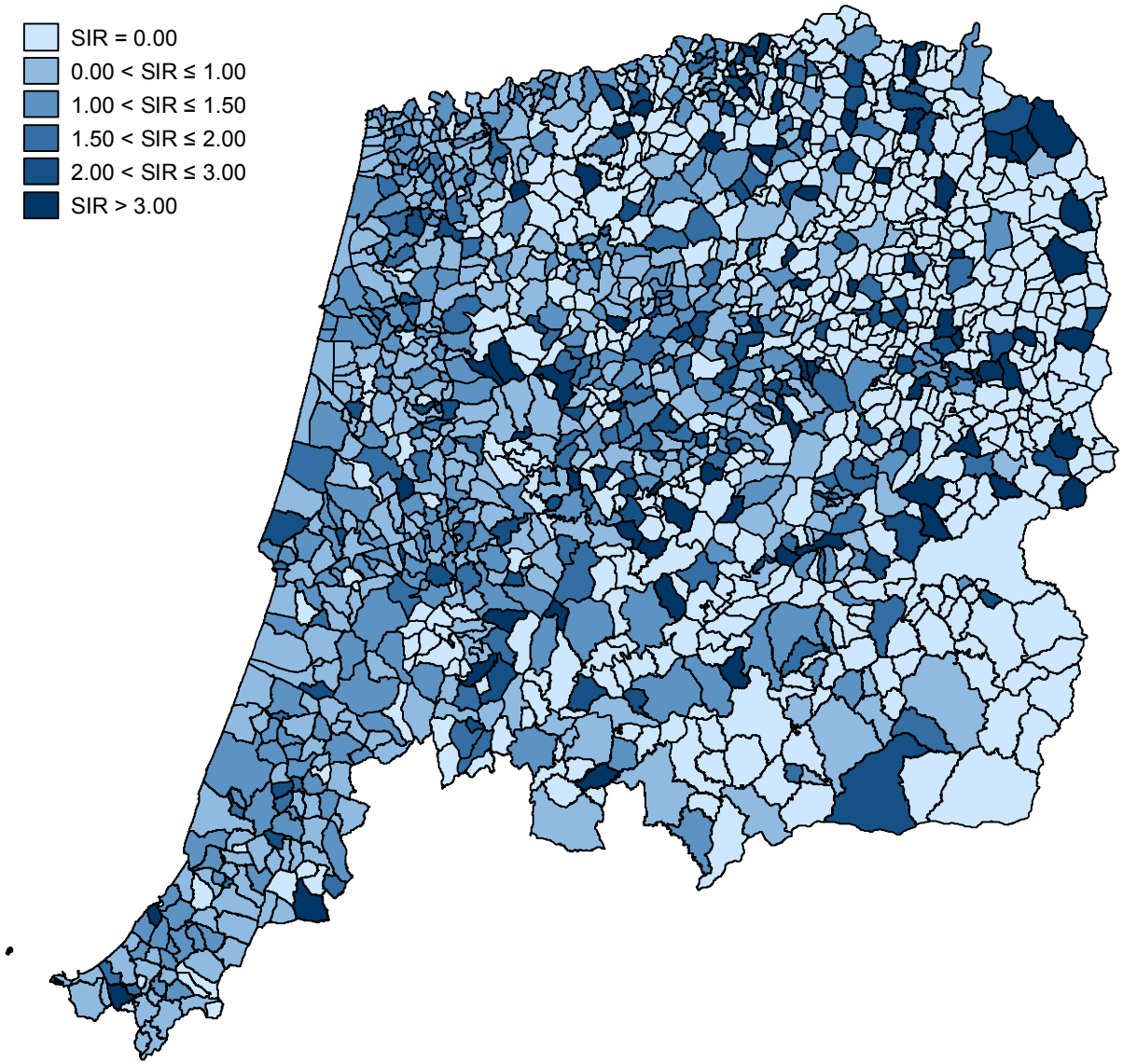
The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
0.00 < SIR ≤ 1.00
1.00 < SIR ≤ 1.50
1.50 < SIR ≤ 2.00
2.00 < SIR ≤ 3.00
SIR > 3.00

Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

The map representing the SIRs by World standardization was reported. No smoothing was applied.

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object cenpmat
-------------------------------------------
        Matrix |                Description
---------------+---------------------------
    Dimensions |                1455 x 1455
     Stored as |                1455 x 1455
 Links         |
         total |                       8416
           min |                          0
          mean |                   5.784192
           max |                         20
-------------------------------------------
warning: spatial-weighting matrix contains 1 island
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by EDI 2011). The EDI 2011 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

Standardization: World

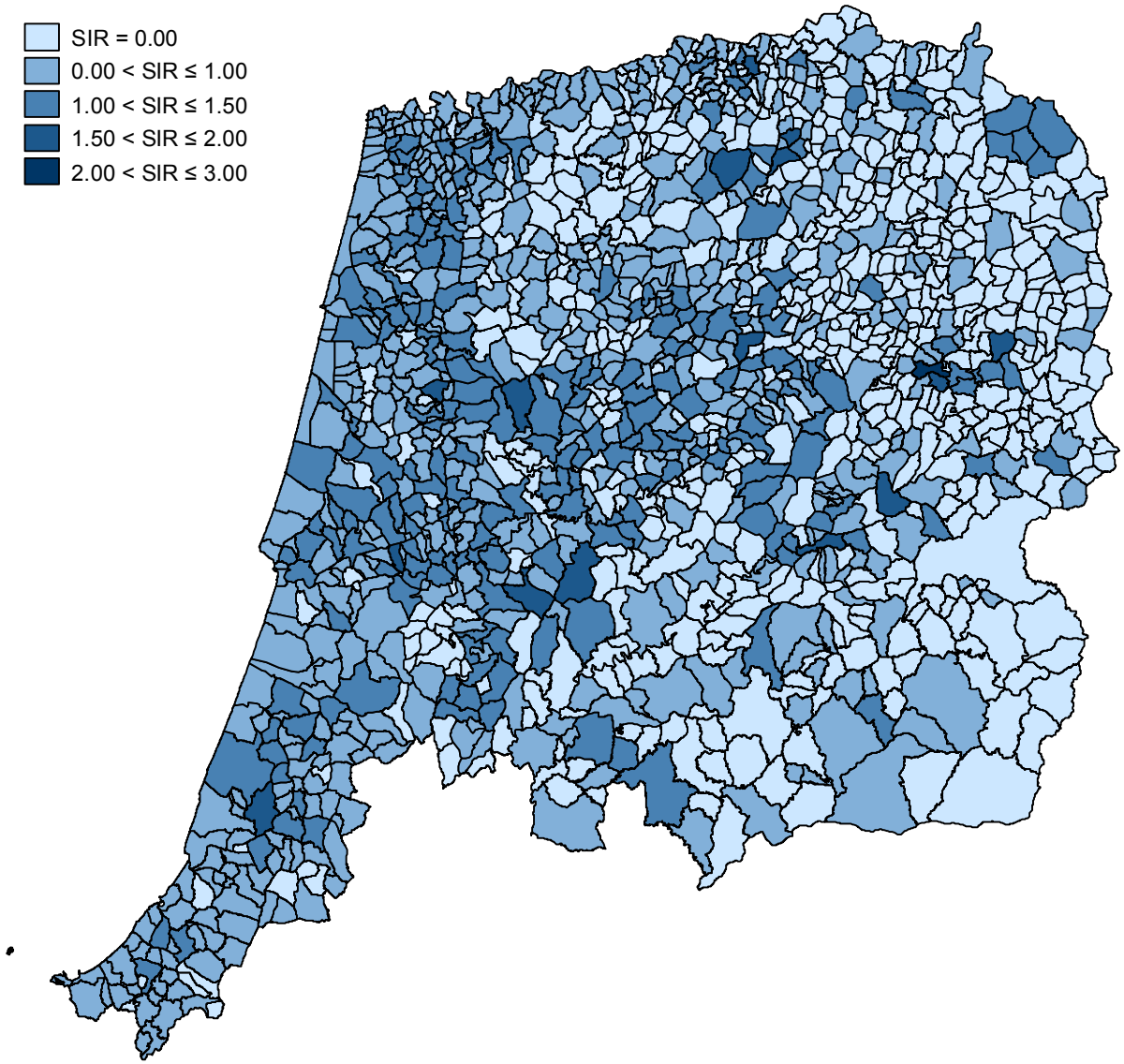
```
Spatial autoregressive model                    Number of obs   =      1455
(GS2SLS estimates)

------------------------------------------------------------------------------
 sir049f_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049f_wor  |
   EDI_SCORE |  -.0100551   .0134862    -0.75   0.456    -.0364875    .0163773
       _cons |   .4310251   .2601003     1.66   0.097     -.078762    .9408122
-------------+----------------------------------------------------------------
lambda       |
       _cons |    1.61889   1.023366     1.58   0.114    -.3868706    3.624651
-------------+----------------------------------------------------------------
rho          |
       _cons |  -.5557108   .8248971    -0.67   0.501     -2.17248    1.061058
------------------------------------------------------------------------------
```

References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

Standardization: World



ANNEX 21 – REPORT WITH ANALYSIS OF BASQUE COUNTRY CANCER REGISTRY DATA

From Basque Country CR data to analysis: first results

Between September and October, the Basque Country CR provided the complete set of data, geo-coded at Census Tract (SU) level, integrating some correction about the right attribution of the 2011 SU.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step-by-step discussion about the best model and software to be used for WASABY's spatial analysis, at Census Tract (named SU for Basque Country) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2005 to the 31st December 2014, for a total number of 3512 female breast cancer patients aged 0-49 year.

The Basque Country CR provided also the 2011 shapefile at SU level for the registry-covered area, the 2011 Census Population and the 2011 local Deprivation Index at SU level.

Merging observed 0-49 cases by Census tract with Basque Country shapefile and 2011 DE dataset

A dataset with SU as key variable was built, merging the information from the Basque Country shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the 2011 DE dataset with the geo-coded Basque Country CR dataset.

The result was a dataset in which the rows reported the SU with the sum of the observed cases (from the Basque Country CR) and all the other information at SU level (from the other datasets).

The 3512 considered incident cases were distributed in 1742 SUs. In 355 SUs where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Spanish (`sir049_esp`), European (`sir049_eu`) and World (`sir049_wor`) population standard were used.

The following table reports the average SIRs for the Basque Country, along with the standard error and the 95% C.I.

	Mean	Std. Err.	[95% Conf.	Interval]
sir049_esp	.955687	.0188055	.9188033	.992570
sir049_eu	1.138251	.0223979	1.094322	1.182181
sir049_wor	1.144715	.0533701	1.039901	1.249528

The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
$0.00 < SIR \leq 1.00$
$1.00 < SIR \leq 1.50$
$1.50 < SIR \leq 2.00$
$2.00 < SIR \leq 3.00$
$SIR > 3.00$

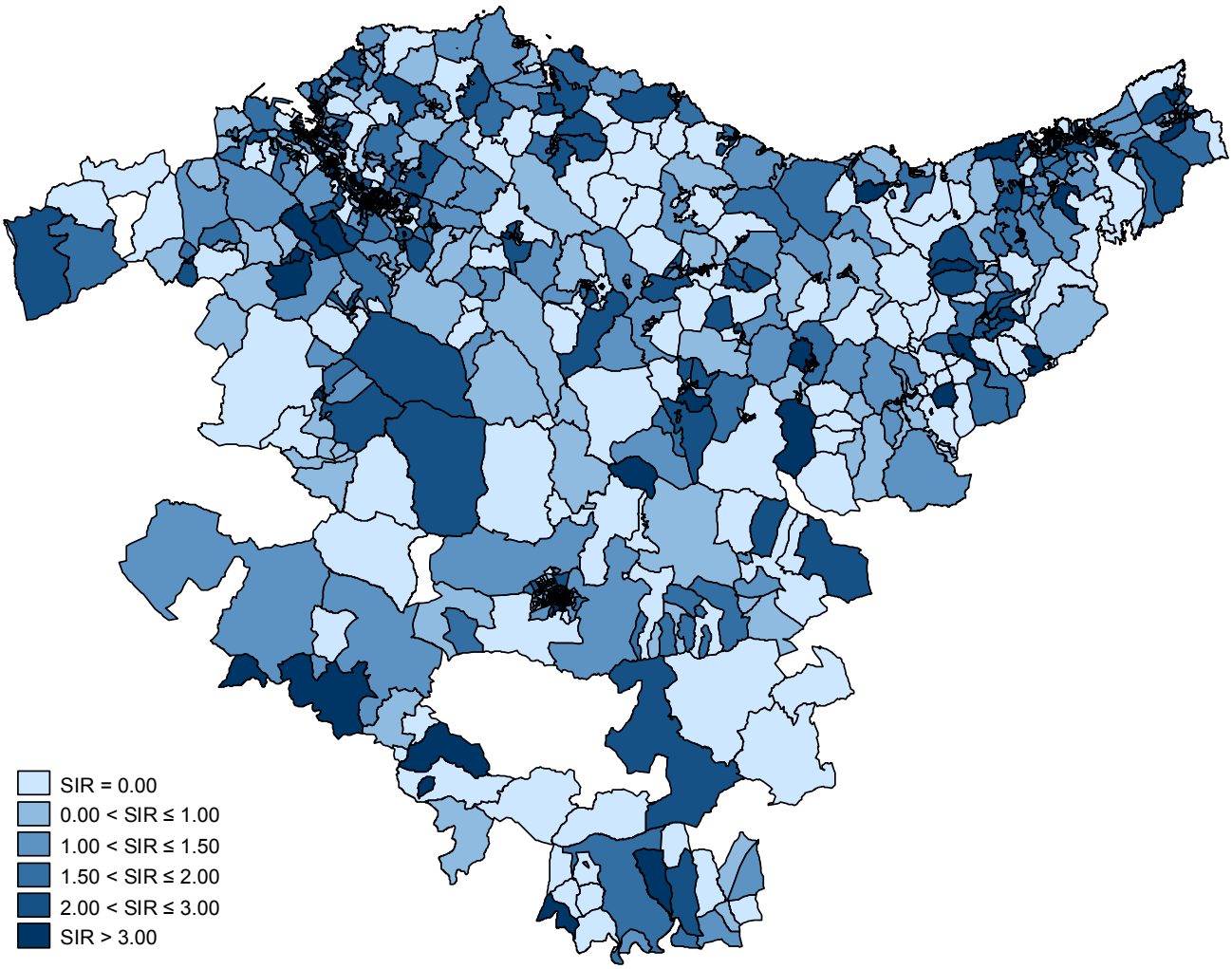
Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by World standardization was reported. No smoothing was applied.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Standard: World



Co-funded by
the Health Programme
of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object basquemat
-----------------------------------------------
        Matrix |                   Description
---------------+-------------------------------
    Dimensions |                   1742 x 1742
     Stored as |                   1742 x 1742
 Links         |
         total |                         10450
           min |                             1
          mean |                      5.998852
           max |                            21
-----------------------------------------------
```

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by DE 2011). The DE 2011 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Spain

```
Spatial autoregressive model                    Number of obs   =      1742
(GS2SLS estimates)

------------------------------------------------------------------------------
  sir049_esp |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049_esp   |
         SES |   .0399235   .1015754     0.39   0.694    -.1591607    .2390076
       _cons |  -.6520059   2.149774    -0.30   0.762    -4.865486    3.561475
-------------+----------------------------------------------------------------
lambda       |
       _cons |   5.795994    7.52625     0.77   0.441    -8.955184    20.54717
-------------+----------------------------------------------------------------
rho          |
       _cons |   -1.37657   .1748912    -7.87   0.000     -1.71935    -1.03379
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =      1742
(GS2SLS estimates)

------------------------------------------------------------------------------
   sir049_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049_eu    |
         SES |     .04755   .1209793     0.39   0.694     -.189565    .2846651
       _cons |  -.7765582   2.560445    -0.30   0.762    -5.794938    4.241821
-------------+----------------------------------------------------------------
lambda       |
       _cons |   5.795994    7.52625     0.77   0.441    -8.955185    20.54717
-------------+----------------------------------------------------------------
rho          |
       _cons |   -1.37657   .1748912    -7.87   0.000     -1.71935    -1.03379
------------------------------------------------------------------------------
```

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Standardization: World

```
Spatial autoregressive model                    Number of obs   =      1742
(GS2SLS estimates)

------------------------------------------------------------------------------
  sir049_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049_wor   |
         SES |   .0589009   .1498588     0.39   0.694    -.2348169    .3526187
       _cons |  -.9619336   3.171659    -0.30   0.762    -7.178271    5.254404
-------------+----------------------------------------------------------------
lambda       |
       _cons |   5.795994   7.526249     0.77   0.441    -8.955183    20.54717
-------------+----------------------------------------------------------------
rho          |
       _cons |   -1.37657   .1748912    -7.87   0.000     -1.71935    -1.03379
------------------------------------------------------------------------------
```

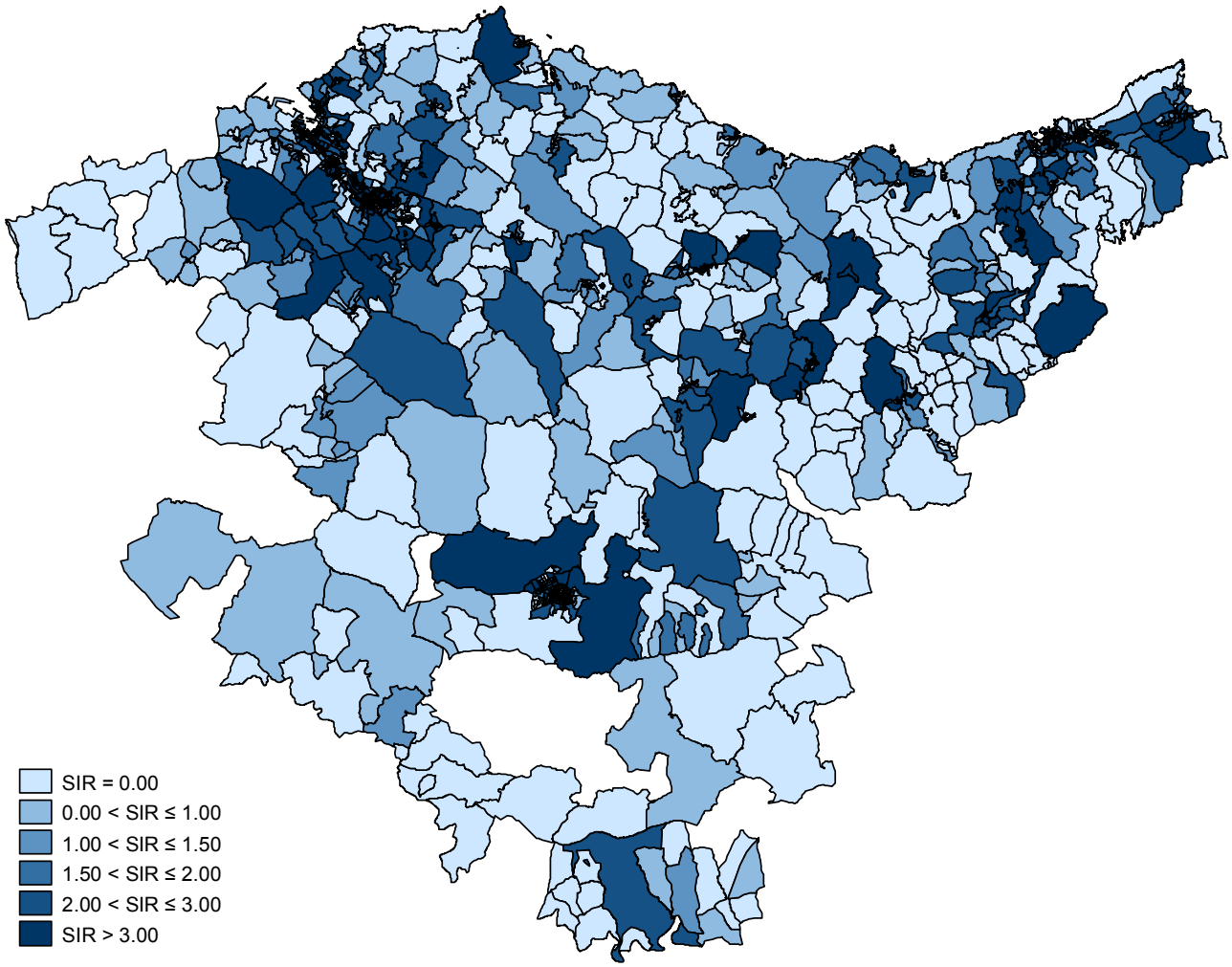
The graphical representation was reported only for the World standardization.

Moreover, also a zoom on the Bilbao, San Sebastian and Vitoria municipalities, the largest and most populous ones of the province, were reported.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

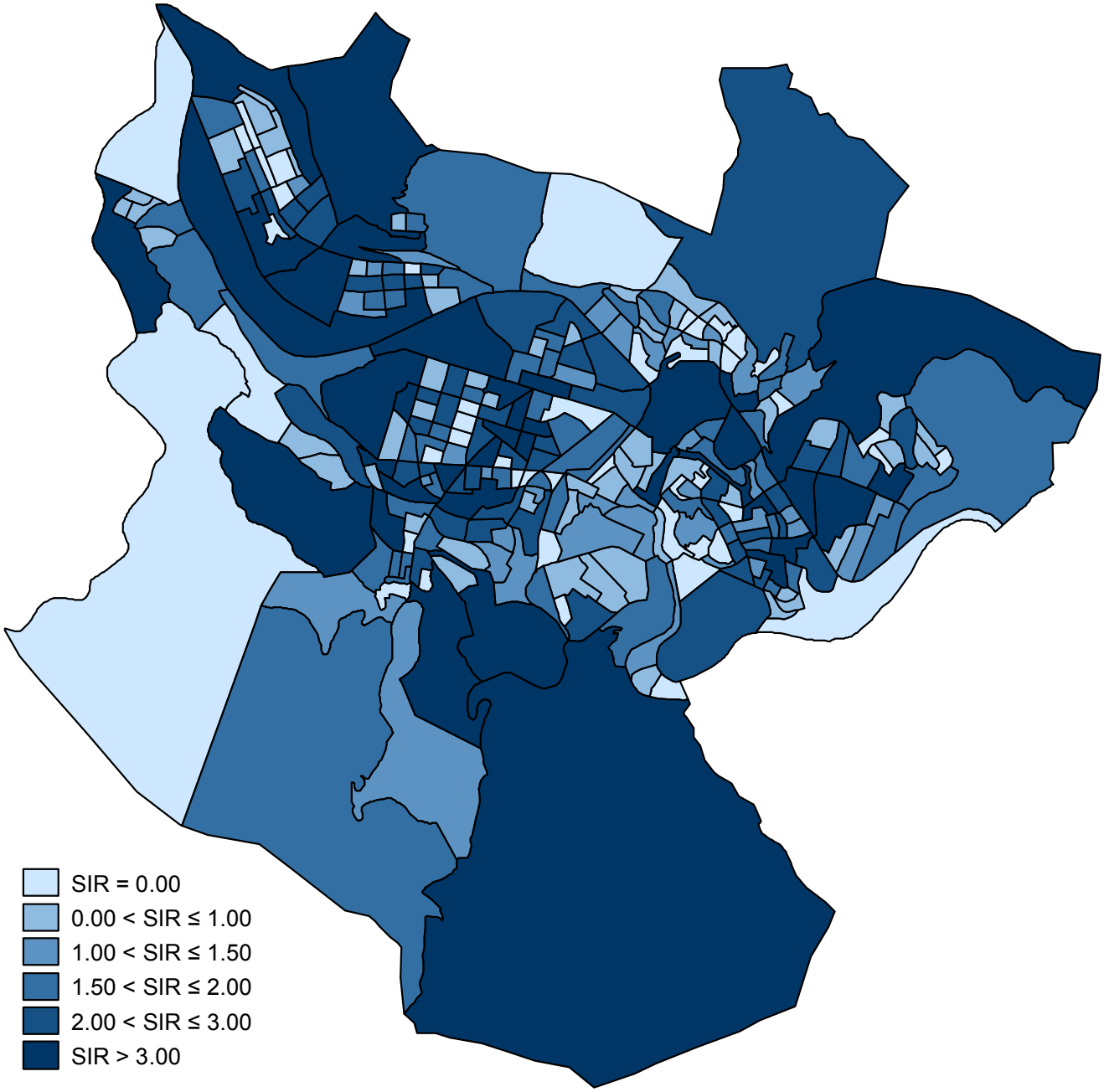
Standardization: World



Co-funded by
the Health Programme
of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Zooming on Bilbao municipality:

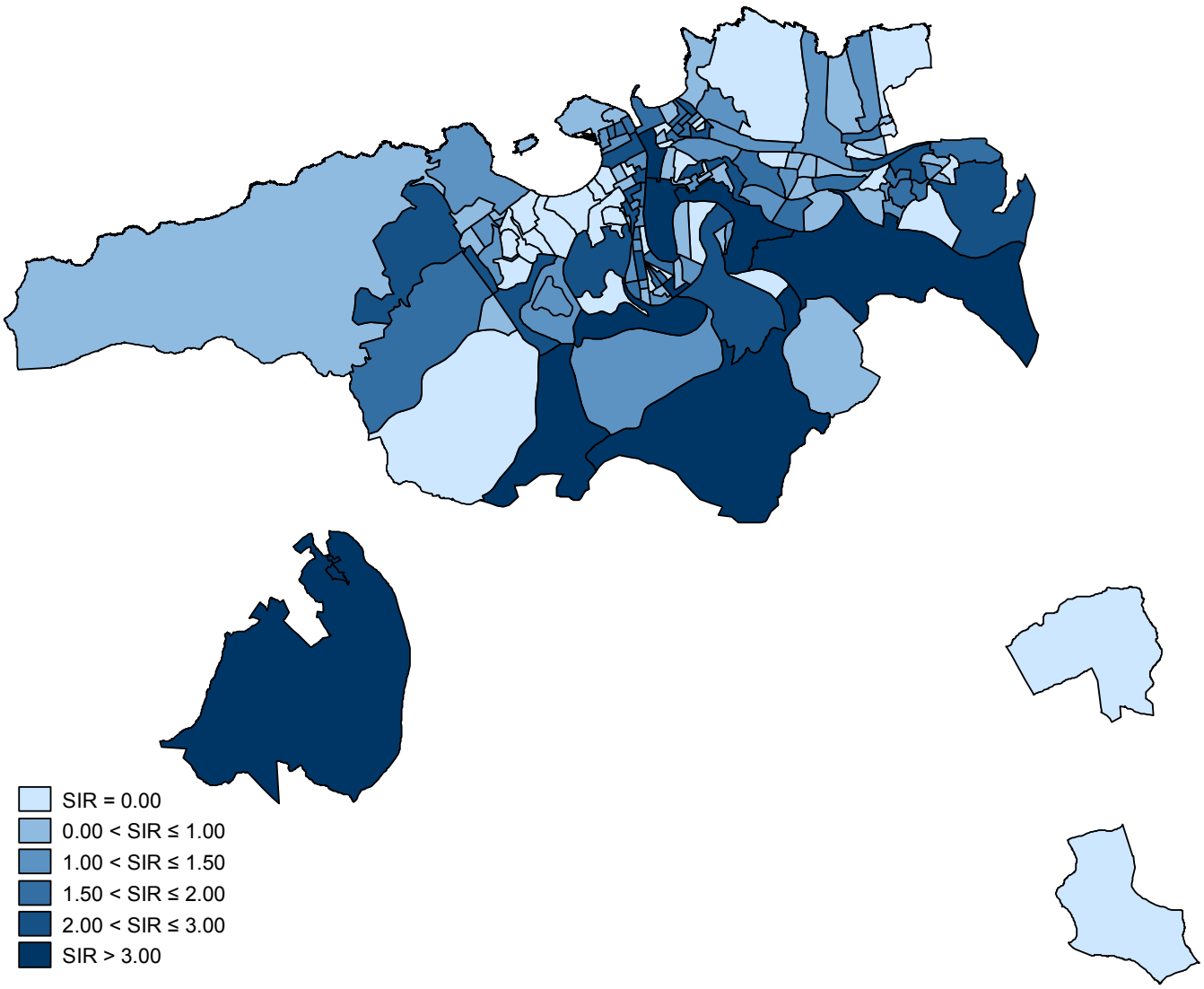


Note: MI_PRINX = 3011 and MI_PRINX = 3295 excluded, because not belonging to Bilbao municipality.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Zooming on San Sebastian municipality:

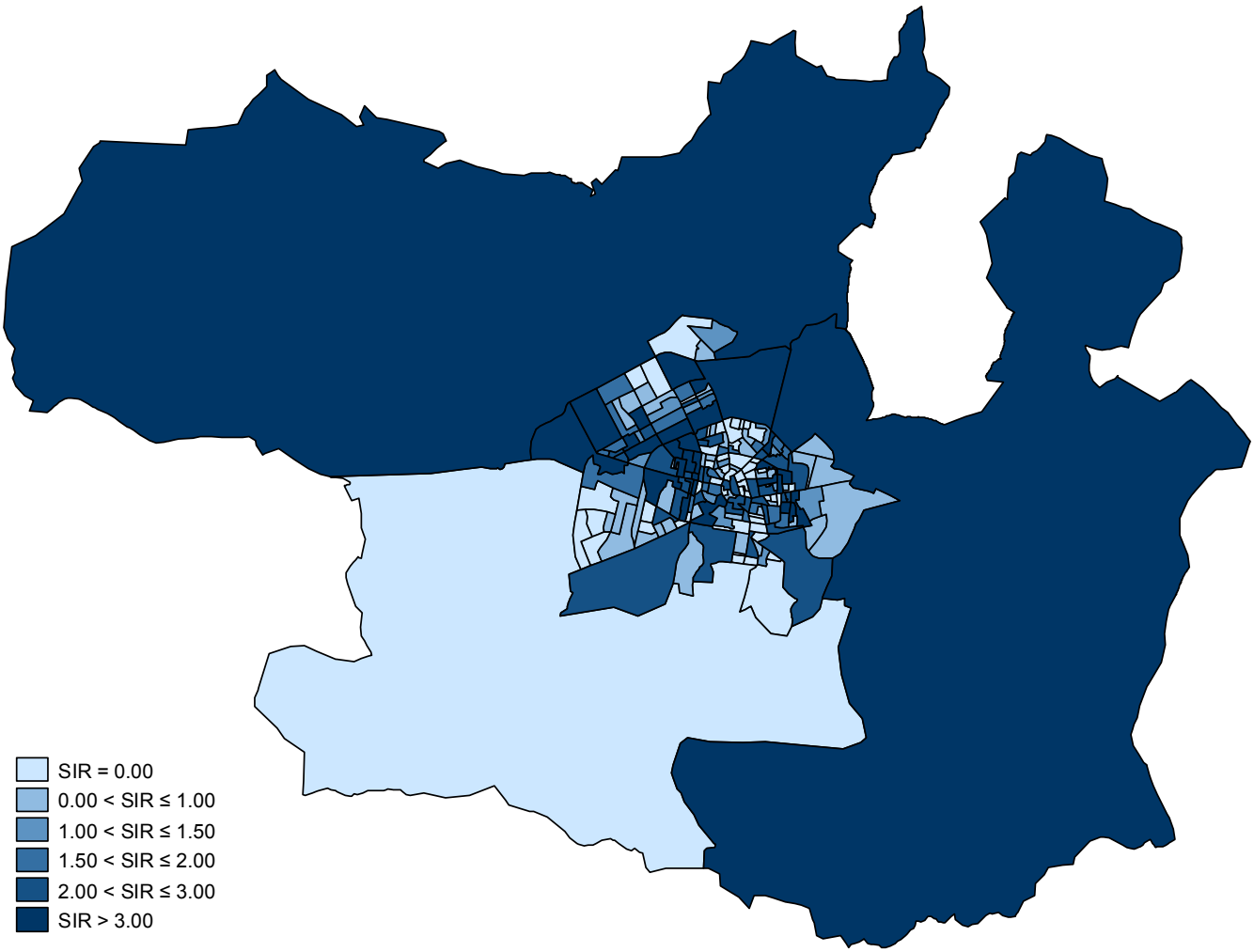


Note: MI_PRINX = 4482, MI_PRINX =4483, MI_PRINX = 4484, MI_PRINX = 4485, MI_PRINX = 4486, MI_PRINX = 4487, MI_PRINX = 4488, MI_PRINX = 4489, MI_PRINX = 4490 and MI_PRINX = 10154 excluded, because not belonging to San Sebastian municipality.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Zooming on Vitoria municipality:



Note: MI_PRINX = 3471, MI_PRINX =3472, MI_PRINX = 3473, MI_PRINX = 3474, MI_PRINX = 3475, MI_PRINX = 3476, MI_PRINX = 3477, MI_PRINX = 3478, MI_PRINX = 3479 and MI_PRINX = 4466 excluded, because not belonging to Vitoria municipality.

References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 22 – REPORT WITH ANALYSIS OF GRANADA CANCER REGISTRY DATA

From Granada CR data to analysis: first results

On the first days of September, the Granada CR provided the second complete set of data, geo-coded at X&Y level.

WASABY's WP4 group chose to test the following spatial analysis technique on this dataset:

- SARAR models by STATA 14 (module: spreg)

While waiting for obtaining STATA 15 version (incorporating BYM models), in the following document the results of the SARAR models, along with some methodological choices and practical actions that allowed the analysis, are presented.

The present document intends to introduce a step by step discussion on the best model and software to be used for WASABY's spatial analysis, at Census Tract (SU2001) level. Corresponding results and a few brief comments for discussion are proposed.

Incident cases were diagnosed from the 1st January 2004 to the 31st December 2013, for a total number of 1245 female breast cancer patients aged 0-49 year.

The Granada CR provided also the 2001 Census Population and the 2001 European Deprivation Index at SU2001 level, while the 2001 shapefile at SU2001 level for the Granada Province was downloaded from the Andalusian Geographic and Statistics Office website.

For 24 cases (1.9% of the total cases) the geo-coded information was missing. They were excluded from the analysis.

Merging observed 0-49 cases by Census tract with Granada province shapefile and EDI 2001 dataset

A dataset with SU2001 as key variable was built, merging the information from the Granada province shapefile (.shp and .dbf dataset from shape file conversion in STATA format) and the EDI 2001 dataset with the geo-coded Granada CR dataset.

The result was a dataset in which the rows reported the SU2001s with the sum of the observed cases (from the Granada CR) and all the other information at SU2001 level (from the other datasets).

The 1221 considered incident cases were distributed in 438 SU2001s. In 168 SU2001s where no cases were present, observed incidence was put to 0 instead of missing value. This choice was due to the need of having no empty cells that could bias the spatial analysis (see Drukker et al., 2013 a, b) and for mapping the area completely.

Result	# of obs.
not matched	168
matched	438
total	606

After this action, all 606 SU2001s of the Granada province were allowed to be used in the SARAR models.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Computing the expected cases and the SIRs by CTs

The expected cases and the SIRs were computed by direct standardization. Both Spanish (`sir049_esp`), European (`sir049_eu`) and World (`sir049_wor`) population standard were used.

The following table reports the average SIRs for the Granada province, along with the standard error and the 95% C.I.

	Mean	Std. Err.	[95% Conf.	Interval]
sir049_esp	.7830136	.0363585	.7116093	.8544179
sir049_eu	.9343286	.0433847	.8491257	1.019532
sir049_wor	1.185828	.0550629	1.07769	1.293965

The SIRs coming from the observed and expected cases were represented by choropleth maps, considering six SIR groups:

SIR = 0.00
$0.00 < SIR \leq 1.00$
$1.00 < SIR \leq 1.50$
$1.50 < SIR \leq 2.00$
$2.00 < SIR \leq 3.00$
$SIR > 3.00$

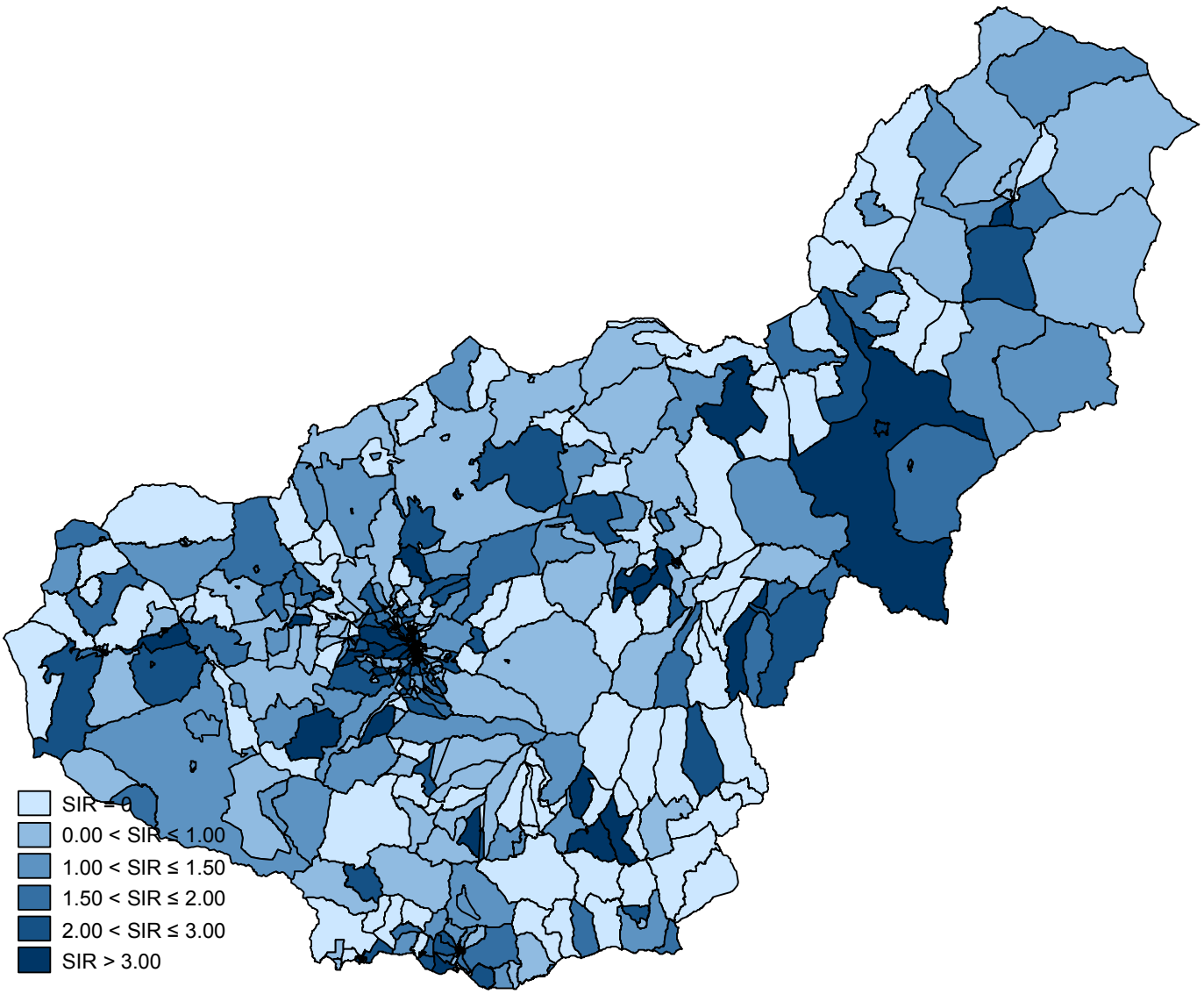
Obviously, these cut-off values were defined only as a possible example. Different categories could be considered.

Only the map representing the SIRs by World standardization was reported. No smoothing was applied.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Standard: World



Building of a space-weighted matrix

The space-weighted matrix were computed in order to be used in the SARAR model for modeling interactions between spatial units (Drukker et al., 2013, b).

```
Summary of spatial-weighting object granamat
------------------------------------------------
        Matrix |                  Description
---------------+--------------------------------
    Dimensions |                    606 x 606
     Stored as |                    606 x 606
 Links         |
         total |                         3486
           min |                            1
          mean |                     5.752475
           max |                           19
------------------------------------------------
```

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Bayesian estimation with SARAR model (spatial-autoregressive model with SAR disturbances) and post-estimation

Finally, SIR estimation with SARAR models was performed, using a generalized spatial two-stage least-squares (GS2SLS) estimator (Drukker et al., 2013, a), considering also the effects of the socio-economic deprivation (by EDI 2001). The EDI 2001 was used in its quantitative format and as an exogenous variable (Drukker et al., 2013, a)

The models were computed for the three standardizations.

Standardization: Spain

```
Spatial autoregressive model                    Number of obs   =       606
(GS2SLS estimates)

------------------------------------------------------------------------------
 sir049f_esp |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049f_esp  |
     EDI2001 |   .0179768   .0091816     1.96   0.050    -.0000187    .0359724
       _cons |  -.2394521   .1615056    -1.48   0.138    -.5559972     .077093
-------------+----------------------------------------------------------------
lambda       |
       _cons |   3.705674   .5489546     6.75   0.000     2.629743    4.781606
-------------+----------------------------------------------------------------
rho          |
       _cons |  -2.933967   .2437493   -12.04   0.000    -3.411707   -2.456227
------------------------------------------------------------------------------
```

Standardization: Europe

```
Spatial autoregressive model                    Number of obs   =       606
(GS2SLS estimates)

------------------------------------------------------------------------------
  sir049f_eu |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049f_eu   |
     EDI2001 |   .0214508   .0109559     1.96   0.050    -.0000223    .0429239
       _cons |  -.2857256    .192716    -1.48   0.138    -.6634421    .0919909
-------------+----------------------------------------------------------------
lambda       |
       _cons |   3.705674   .5489546     6.75   0.000     2.629743    4.781606
-------------+----------------------------------------------------------------
rho          |
       _cons |  -2.933967   .2437493   -12.04   0.000    -3.411707   -2.456227
------------------------------------------------------------------------------
```

Standardization: World

```
Spatial autoregressive model                    Number of obs   =       606
(GS2SLS estimates)

------------------------------------------------------------------------------
 sir049f_wor |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sir049f_wor  |
     EDI2001 |   .0272248   .0139049     1.96   0.050    -.0000284     .054478
       _cons |  -.3626361   .2445907    -1.48   0.138     -.842025    .1167527
-------------+----------------------------------------------------------------
lambda       |
       _cons |   3.705674   .5489546     6.75   0.000     2.629743    4.781606
-------------+----------------------------------------------------------------
rho          |
       _cons |  -2.933967   .2437493   -12.04   0.000    -3.411707   -2.456227
------------------------------------------------------------------------------
```

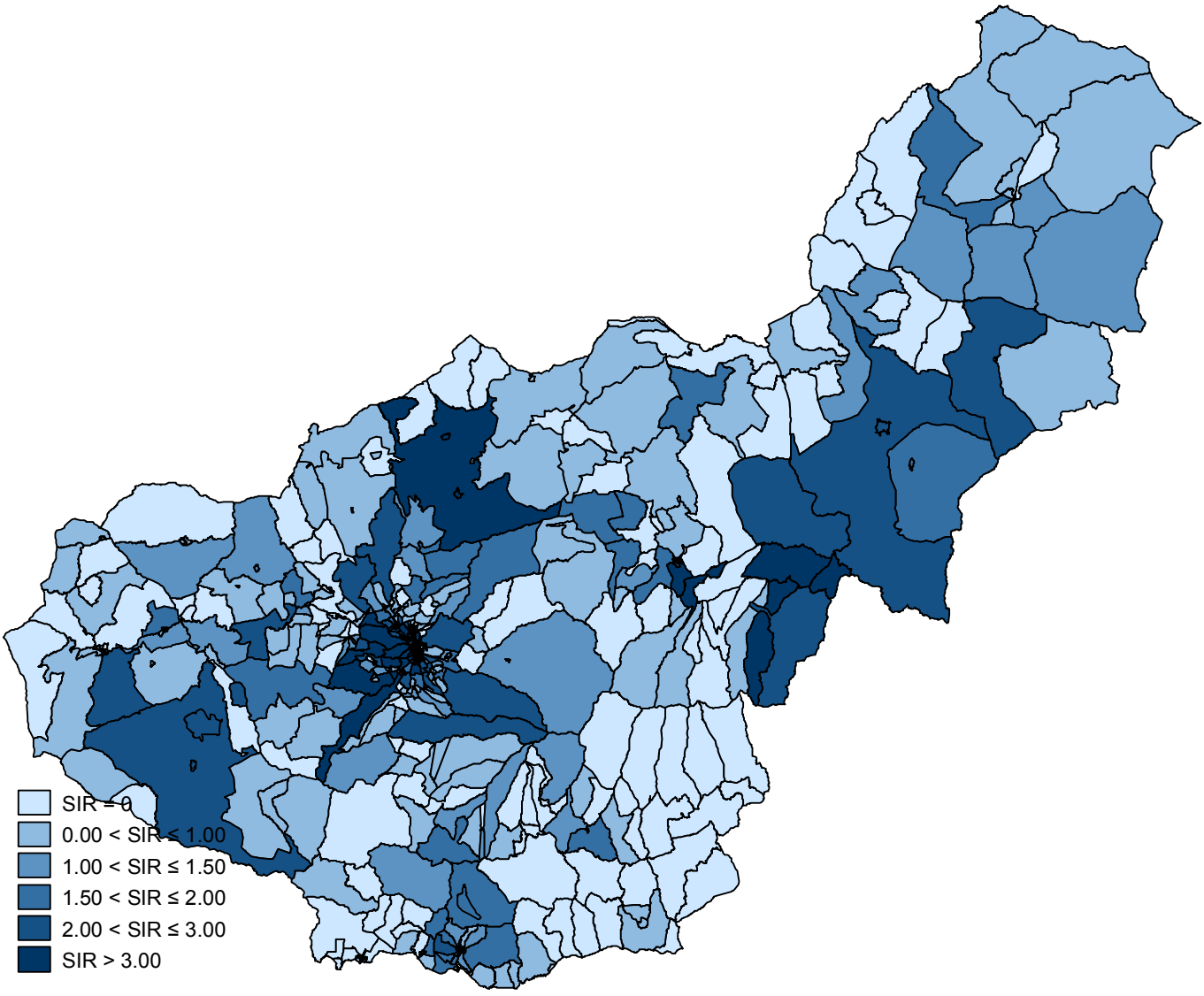
The graphical representation was reported only for the World standardization.

Moreover, also a zoom on the Granada municipality, the largest and most populous of the province, was reported.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

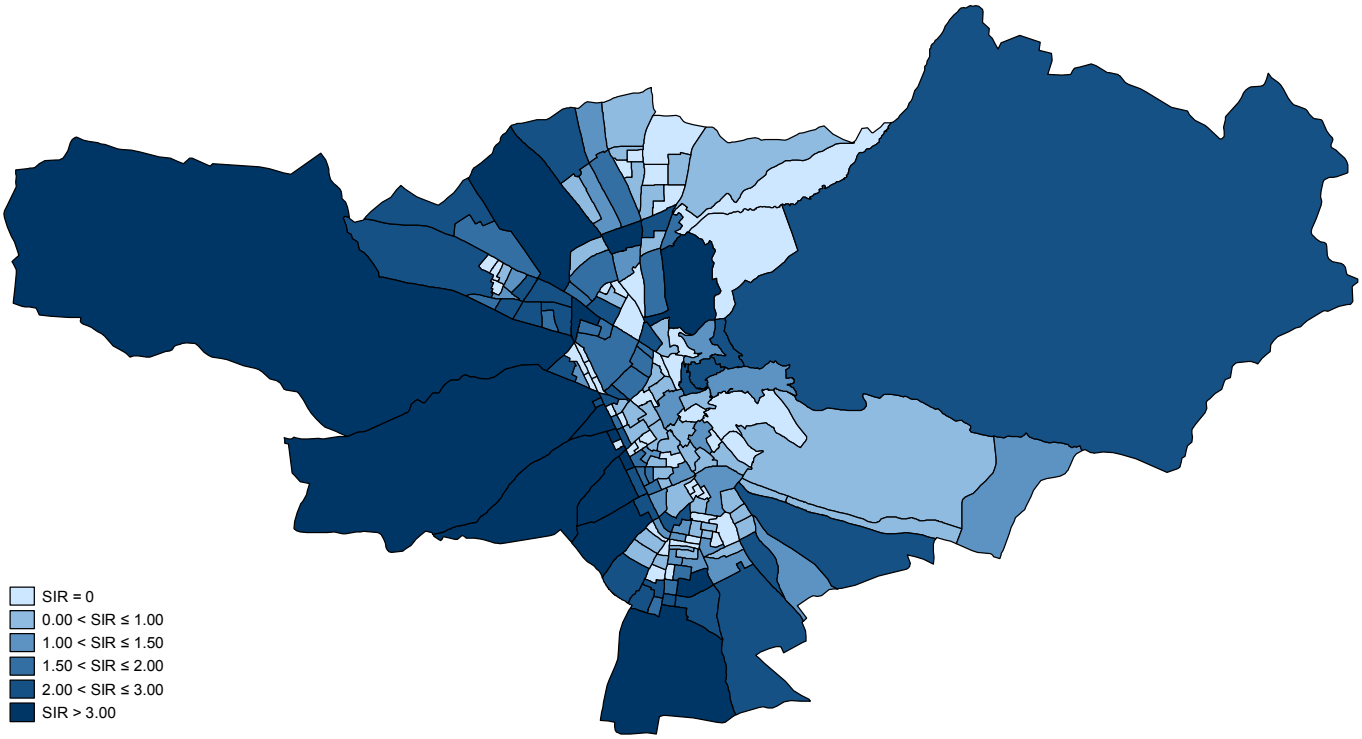
Standardization: World



Co-funded by
the Health Programme
of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Zooming on Granada municipality:



References

Drukker DM, Prucha IR, Raciborski R. (2013) (b) Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal; 13(2): 221-242.

Drukker DM, Peng H, Prucha IR. (2013) (b) Creating and managing spatial-weighting matrices with the spmat command. The Stata Journal; 13(2): 242-286.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Introduction and general objective of WP7

The WASABY project concerns the design of a model able to identify areas with higher cancer rates, in order to study whether the contamination of the most persistent pollutants in soil and in deep water can cause an increase in the risk of breast cancer. One of the main objectives of WP7 is to analyze the relationship between environmental data and the incidence of breast cancer following the following aim: identify, through a **scientific literature** review, persistent contaminants in the aqueous matrix (especially in deep waters) and in the soil matrix, , assessing environmental risk factors and exposure to contaminants, in relation to breast cancer. Several scientific articles were found to correlate exposure to persistent environmental contaminants, mainly, but not only, with POPs (Persistent Organic Pollutants), present in the two environmental matrices (water and soil) and the risk of breast cancer for women.

Methods and means

When we searched scientific articles correlating exposure to Persistent Organic Pollutants (POPs) and the risk of having breast cancer for women, we considered and selected the most important articles of population studies (e.g., case-control and cohort studies) starting from scientific reviews, meta-analysis and IARC monographs on POPs exposure through biological matrices before or after breast cancer diagnosis (e.g., blood samples, plasma and breast tissue) . In our literature research, we included ‘‘breast cancer’’ in combination with environmental pollutant; criteria include original studies published in a peer-reviewed journal with a case-control, nested case-control or cohort design, estimating breast cancer risk (odds ratio, risk ratio or hazard ratio) associated with environmental exposure to combustion products, pesticide, drinking water, organic solvent, heavy metals (cadmium, lead, nickel, etc.), polychlorinated biphenyl (PCBs), organochlorine pesticides, DDT,p,p’; DDT,o,p’; DDD,p,p’; DDE,p,p’; Hexachlorobenzene (HCB), Dieldrin, Alachlor, polycyclic aromatic hydrocarbon (PAH), triazine and metabolites, Perfluoroalkylated substances (PFAS), cadmium (heavy metal), and trihalomethanes (THMs) (bromoform, bromo-dichloro-methane, dibromochloromethane, chloroform). In our search, we also incorporated ‘‘breast cancer’’ and each of the chemicals identified as a mammary carcinogen. We limited inclusion to articles published from the late 90s to the present day, for all contaminants, because there are different useful reviews covering earlier articles and we sought to update these reviews, minimizing redundancy. Articles were reviewed to identify and evaluate inclusion criteria for study participants, comparability of control or reference groups, exposure measurement method, control for confounding, and the strength of observed associations. All studies included are controlled for age and sex, and are all restricted to females. A critical review for each study was entered in a database. Results were summarized in tables for the most-studied exposure sources in Annex 1-8.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Environmental contaminants considered

In this part of the WP7 we set out to review and synthesize epidemiologic evidence concerning breast cancer and environmental pollutants identified as mammary carcinogens or endocrine disrupting compounds, including: persistent organochlorine pollutants (POPs) and other contaminants routinely monitored by the environmental agencies of Europe in the various environmental water and soil matrices, and included in specific standardized databases.

They are environmentally persistent and lipophilic. They are frequently detected in food, soil, and dust, concentrate up the food chain, and are found in human breast milk and adipose tissue. Residues can be readily measured in blood and breast tissue, providing a way to quantify exposure, although these measures are invasive and expensive; therefore, as a practical matter, levels cannot be measured repeatedly in an individual. They include:

3.1 Polychlorinated biphenyls (PCBs). PCBs are a group of man-made organic chemicals consisting of carbon, hydrogen and chlorine atoms.

3.2 Dichlorodiphenyltrichloroethane (DDT) and its metabolites (DDD, DDE) and others organochlorine compounds as 1,1-Dichloro-2,2-bis(4-chlorophenyl)ethene (DDE) and dichlorodiphenyldichloroethane (DDD) and pesticides like Dieldrin, hexachlorobenzene (HCB), Hexachlorocyclohexane, Chlordane.

3.3 Dioxins are unwanted by products of chemical processes that contain chlorine and hydrocarbons. There are at least 100 different kinds of dioxins such as 2,3,7,8-tetrachlorodibenzo-pdioxin (TCDD).

Others pollutants

3.4 Polycyclic aromatic hydrocarbons (PAHs). Ubiquitous group of several hundred chemically-related, environmentally persistent organic compounds of various structures and varied toxicity

3.5 Perfluoroalkylated substances (PFAs). A large group of man-made chemicals. PFAS include

- perfluoro carboxylated acids (PFCAs) perfluorosulfonated acids (PFSAs), which include
- perfluorooctanoic acid (PFOA) and perfluorooctane sulfonate (PFOS)

3.6 Triazine group of chemicals are most widely used herbicides in the world. This group consists of atrazine, simazine, propazine and cyanazine.

3.7 Heavy metals. In principle Cadmium can play an important role in the risk of breast cancer

3.8 Trialomethanes. Trihalomethanes (THMs) are halogen-substituted single-carbon compounds most commonly present in drinking-water are chloroform ($CHCl_3$), bromodichloromethane or dichlorobromomethane ($CHBrCl_2$) (BDCM), dibromochloromethane or chloro-dibromomethane ($CHClBr_2$) (DBCM), and bromoform ($CHBr_3$).

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

The following chapters of the Report are all structured in the same way:

In the **Contaminants considered**, we list the main properties of the family of contaminants considered, from the molecular structure, to the chemical-physical properties, to the persistence in the environment and to the main contamination in the various environmental matrices, especially in water and soil.

In the **Source,** the main sources of anthropogenic emissions of the contaminants are considered in the various environmental matrices (air, water, soil, sediments and biota) and the main use.

In the **Mean of exposure,** the different modalities of exposure of the contaminants considered towards the man are evaluated, passing for example through the trophic chain, the contact, the respiration also evaluating the potential toxic effects of the contaminants.

In the **Regulation and/or persistence,** the various directives and laws that regulate the use of the contaminants considered are evaluated.

In the **Evidence on (breast) cancer risk**, scientific studies of population (case control / cohort) are presented where a correlation was found between the different persistent contaminants considered and breast cancer.

In the **IARC Monography**, the most important scientific studies on the population that assessed the exposure to the considered persistent contaminants (also family of contaminants) and the correlation with the breast cancer are often analyzed through a biological sampling of women cases and controls considered (blood samples and breast tissue).

In the **Reviews and meta-analyses,** we analyzed the main reviews and meta-analyses in literature that covered the exposure to the considered persistent contaminants in correlation with breast cancer.

In the **ANNEX,** results are summarized as **Annexes 1-8** in form of tables for the main classes of persistent contaminants. The tables all have the same approach: the name of the first three researchers of the scientific paper; the design of the observational study (case-control, cohort, etc.); the geographic location; exposure to the main persistent contaminants; the years of interest of the study; the relative risk observed; the main comments relevant to the scientific article; the possible covariates and methods used.

Co-funded by
the Health Programme
of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.1. Polychlorinated biphenyls (PCBs)

Contaminants considered

PCBs are a class of organic compounds with one to ten chlorine atoms attached to a biphenyl, which is a molecule composed of two benzene rings. Polychlorinated biphenyls (PCBs) represent a class of 209 congeners, classified based on the number and relative position of the chlorine atoms on the biphenyls. PCBs are highly persistent in the environment and today vast areas near former production sites are still polluted. PCBs contamination results particularly alarming when impacting agricultural soil, since they effectively enter the food chain. PCBs concentrate in soil and sediment but can also adsorb to small organic and mineral particles, therefore traveling far from the initial contamination site, carried by wind and water. Because they are so persistent and lipophilic, PCBs tend to accumulate in animals and human tissues.

Source

PCBs are industrial chemicals that have been used for various commercial applications such as hydraulic fluids, printing inks or capacitor dielectric liquids. Transformers probably represent the largest source of PCB and are easily identifiable since their nameplates indicate that they are insulated with PCB dielectric. Other PCB applications include plasticizers in paints and cements, stabilizing additives in flexible PVC coatings of electrical cables and electronic components, pesticides extenders, cutting oils and lubricants, flame retardants, such as adhesives, wood floor finishes, paints, in de-dusting agents, in waterproofing compounds.

Mean of exposure

Once in the environment, the PCBs enter the food chain: most of human exposure to PCB is through food. Such exposure also has neurotoxic and immunotoxic effects. In addition, PCB undergo long-range transport through air, water and migratory species. They travel across international boundaries and are deposited far from their point of release, and accumulate in terrestrial and aquatic ecosystems. There is therefore an urgent need for PCB to be inventoried, taken out of use and managed in an environmentally sound manner. Adults are mainly exposed through the consumption of dairy products, meat and fish. PCBs exist in 12 of these 209 molecules (which are coplanar) are characterized by a toxicity compared to that of dioxins and are defined dioxine-like PCBs. These have serious effects on human's health, such as immune and endocrine system disruption, interference with fetal development, acute and chronic toxicity, carcinogenicity.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Regulation and/or persistence

The directive 96/59/EC establishes the requirements for an environmentally friendly disposal of PCBs, aiming to dispose of all PCBs contained in them; in addition, Member States must take an inventory of large equipment containing PCBs, must adopt a plan for the disposal of inventoried equipment and lines for the collection and disposal of non-inventoried equipment (small electrical equipment very often found in appliances manufactured before the ban on marketing of PCBs).

Experience in many developing countries has shown that the lack of an adequate regulatory framework on the PCB can seriously hamper the inventory process. In many cases, roles and responsibilities are not clear, which means that no institution is responsible for regulating and monitoring the use of PCBs.

Evidence on (breast) cancer risk

Studies on workers' PCBs have shown that they can be associated with certain types of cancer in humans, such as liver and biliary cancer. As a result, the International Agency for Research on Cancer IARC has classified the PCB as a "human carcinogen" group 1. Regarding the Evidence on (breast) cancer risk on the exposure of PCBs there are several studies of population well studied by the recent IARC Monograph (year 2016) on PCBs, and also in several reviews and meta-analyzes of scientific literature.

Different case-control studies do not find a statistically significant relationship between exposure to PCBs (through blood samples and adipose tissue analisys) and breast cancer, but some articles, confirmed in the 2016 IARC Monograph and in the most recent reviews of scientific literature and meta-analysis, find a statistically significant association between exposure to some PCBs congeners and breast cancer.

For example, in the hospital-based case-control study conducted in 1994-1996 in Mexico City **[Lopez-Carrillo 1997],** 141 cases aged between 20 and 79 years from participating hospitals and an equal number of controls without cancer were enrolled. Serum levels of PCBs were analysed in 95 histologically confirmed breast cancer cases and 95 hospital controls, 2079 years of age. After adjusting for established risk factors, there was no evidence of a relationship between PCBs and breast cancer risk (OR=1.31, 95% CI 0.33‑5.21) for the high category of exposure. This study lends no support to the case for a role for PCBs in breast cancer aetiology.

A recent case-control study carried out in Alaska **[Holmes 2014]** analyzed blood and adipose tissues collected after the diagnosis of breast cancer; also in this case they did not observe a significant association between some PCB congeners (PCB 138, 158, 153, 180) with risk of breast cancer.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

The purpose of a case-control study **in Spain [Lucena 2001]** was to examine the possible relationship between different PCBs congeners and breast carcinomas. In this study all the women treated with excision biopsy due to the lump were included, a questionnaire was performed for all the cases and controls, the body mass index was computed, a histopathological examination of the mass removed and the concentration of PCB levels in breast fat was calculated through samples of adipose tissue of which 48.5% were from women with benign lesions and 51.5% with malignant lesions. In multivariate analysis, the PCB-28 congener was found to be the most important risk factor (OR=9.6, 95% CI 3.8-24.4). The variables associated with the malignant lesions in the univariate analysis were age, lactation period, overweight, n-28 PCB and other PCB congeners. Other risk factors have been identified as age, alcohol, low parity and overweight.

A case-control study done in **Tunisia [Arrebola 2015]** finds an association (again from the blood samples analyzed after diagnosis) between the two PCB congeners (PCB138 and 180) in 69 cases of breast cancer and 56 controls. PCB138 and PCB180 were positively associated with breast cancer risk but only in univariate analysis.

In another recent case-control study **[Wielsoe 2017]** the Greenland inuit women were recruited in the period 2000-2003 and in the period 2011-2014. The diagnosis of breast cancer was confirmed by the histology. They found weak but positive associations between high serum concentrations (average/maximum vs. lower tertile) on the individual congeners of the more lipophilic PCBs (PCB99, PCB138, PCB153, PCB170, PCB170 and PCB183) and on the total PCBs there is an association with the risk of breast cancer.

IARC Monography

The International Agency for Research on Cancer (IARC) in the Monograph on the evaluation of carcinogenic risk to humans “Polychlorinated Biphenyls and Polybrominated Biphenyls” Vol. 107 (2016) **[IARC 2016]** classified PCBs as carcinogenic to humans based on evidence of malignant melanoma and positive associations with breast cancer and non-Hodgkin lymphoma. Some congeners PCB-118 and PCB-126 and highly chlorinated commercial mixtures (Araclor 1260, Araclor 1254, and Kanechlor 500) exhibit dioxin-like activity and bind to the aryl hydrocarbon receptor (AhR).

IARC’s evaluation of the carcinogenicity and exposure of PCBs and breast cancer was considered in many scientific articles.

The relationship between serum concentration of PCBs and breast cancer has been evaluated in different case-control studies.

In the first five USA case control studies that we have analyzed, we have found a statistically significant relationship between some PCBs congeners and breast cancer.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

In a case-control study performed in New York **[Moysich 1998]**, 154 women with postmenopausal incidental breast cancer and 192 controls of the same age were compared by performing blood test samples, detecting specific chemical analyzes by serum concentrations of 73 PCBs congeners. No association between total PCBs exposure was found, but an increased risk was evident for the less chlorinated PCBs (OR=1.66, 95% CI 1.07-2.88 for the second and third tertile) among Parisian women who had never breast-fed; the magnitude of the risk was higher in association with total PCBs (OR=2.87, 95% CI 1.01-7.29) and moderately chlorinated PCBs (OR=3.57, 95% CI 1.10-8.60).
In a case-control study conducted between 1993 and 1996 in North Carolina, USA, **[Millikan 2000]** breast cancer cases (778) and controls (659) among African American and white women were classified by age, race, menopausal status, BMI, breastfeeding, therapy hormone associated replacement and income. Total PCBs concentrations in plasma lipids were measured. The risk was particularly high for African American women with BMI>34.2 (total third of tertiary PCBs, OR=4.92, 95% CI 1.63-14.83); there was no risk among white women.

In another case-control study in Connecticut, USA **[Falck 1992]** mean concentrations of PCBs in the breast tissue of 20 women with breast cancer were significantly higher (p=0.02) than in 20 women with benign breast disease, and the association persisted after controlling for age, smoking, and BMI. There was approximately 1% increased risk for every 10 ppb of PCBs in adipose tissue.

In a case-control study in Connecticut, USA **[Holford 2000]** calculated the risk of breast cancer in relation to increases in exposure to 10 ng/g at different PCBs congeners. The researchers found that the PCB-180 and the PCB-183 congeners were associated with a statistically significant increase in breast cancer risk, for example for PCB-183 (OR=1.82, 95% CI 1.12-2.98).

An interesting case-control study that included primarily young women (premenopausal women) in Oakland, California [**Cohn 2012]** compared serum concentrations of 16 PCBs in serum samples stored at the beginning of postpartum collected between 1959 and 1967 by 112 breast cancer cases and 112 age-matched checks. The median time from blood sampling to diagnosis was 17 years and the average age at diagnosis was 43 years. No association between breast cancer risk and total PCB counts was reported, but for the PCB-203 congener a statistically significant association was found with an increased risk (OR=6.3, 95% CI 1.9-21.7) for the highest compared to the lower quartile.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Also in a case–control study conducted in 1994–1997 in Quebec City, Canada **[Demers 2002]**, plasma concentrations of fourteen PCBs congeners were measured in 314 women with cancer of the breast and 523 controls (219 hospital controls, 304 population controls). The results suggest that exposure to dioxin-like PCBs increases breast cancer risk, specially associated with a total concentration of the three mono-ortho-substituted congeners of PCBs 105, 118, and 156 expressed as 2,3,7,8 -tetrachlorodibenzo-p-dioxin toxic equivalents (OR=2.02, 95% CI 1.24-3.28, fourth vs. first quartile).

In another case-control study in Mexico **[Recio-Vega 2011]** 70 breast cancer cases were compared with 70 hospital controls by taking blood samples to measure 20 PCBs congeners. An increased risk of breast cancer was evident for total PCB (OR=1.09, 95% CI 1.02-1.16) and for exposure groups 2b (OR=1.90, 95% CI 1.25-2.88) and group 3 (OR=1.81, 95% CI 1.08-3.04) and group 4 (OR=1.57, 95% CI 1.20-2.07), defined according to the PCBs classification of **[Wolff 1995]**.

In a large case-control study on the population of environmental exposures to PCBs and breast cancer conducted in 1996-1997 in Long Island, NY, USA, **[Gammon 2002]** serum concentrations of 24 congeners PCBs were measured for 646 cases and 429 controls, with results presented for the four most common congeners (PCB-118, PCB-138, PCB-153 and PCB-180). There was no association between breast cancer and the concentration of the sum of the four PCBs; no effect of lactation, state of menopause, stage of disease or hormone receptor status.

In northern Europe, several case-control studies on breast cancer and exposure to PCBs, always measured by blood tests and/or through adipose tissue, have been conducted, for example in Sweden **[Liljegren 1998]**. The concentrations of PCBs were measured in mammary adipose tissue of 43 women with breast cancer and 35 controls. Equity-adjusted odds ratios for age and parity did not show any association with total PCBs congeners in all subjects.

In the Copenhagen City Heart Study (Denmark) **[Hoyer 1998]** a nested case-control study was conducted. Serum samples were obtained in 1976 from a cohort of 7712 women aged 20 years or older: after excluding subjects without a valid serum sample, 240 cases and 447 controls were included in the study. Concentrations of 28 PCBs congeners were detected in serum. No association between the risk of breast cancer and the correct lipid concentrations of the sum of PCBs or specific congeners has been reported.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Five years later, the same subjects in the study performed a second serum sample **[Hoyer 2000]**. The analyses were conducted in this group for four congeners of common PCBs. A statistically significant risk was found for subjects in the highest quartile of the PCB congener concentration 138 (OR=2.1, 95% CI 1.0-4.4) while for total PCBs (OR=1.6) and for the congeners PCB-118 (OR=1.9) and PCB-153 (OR=1.3) the association was not significant.
Within the same cohort **[Hoyer 2001]** the overall survival of breast cancer in relation to serum organochlorine concentrations based on the state of estrogen receptors was evaluated. In general, the risk of dying among women with the highest organochlorine exposure level was higher among women with ERP than ERN breast cancers, but the only statistical significant relationship was observed for ΣPCB (RR=2.5, I vs. IV quartile, 95% CI 1.1-5.7), but not dose-response relation was apparent.
Two case-control studies were conducted in Belgium [**Charlier 2003 and 2004]**: in the first study, on 100 cases of breast cancer and 100 surgical controls, the concentrations of PCB-101 and PCB-153 congeners were significantly higher for cases compared to controls; in the second study on 60 cases and 60 controls, an association was reported only for the congener PCB-153 (OR=1.8, 95% CI 1.4-2.5) after adjustment for age and reproductive risk factors .

Also other case-control studies conducted in USA **[Krieger 1994; Dorgan 1999; Helzlsouer 1999; Wolff 2000; Laden 2001; Zhang 2004; Rubin 2006; Gatto 2007]**, in Canada **[Demers 2000]**, in Norway **[Ward 2000]** and in Japan **[Itoh 2009]** found no association between serum PCBs levels and breast cancer risk.

Different case-control studies were conducted on the relationship between the concentration of PCBs in adipose tissue and the risk of breast cancer in the USA **[Zheng 2000a and 2000b; Rusiecki 2004]**, in Denmark [**Raaschou-Nielsen 2005]** and these found no association between adipose tissue PCBs levels and breast cancer risk.
On the contrary, other large case-control studies that have related exposure to total or to different congeners PCBs through the analyzes of adipose tissue in women have found a statistically significant risk with breast cancer.

For example in the case-control study performed by **[Stellman 2000]** in Long Island, New York, USA, concentrations of 14 PCBs congeners in adipose tissue were analyzed in several women with breast cancer and several hospital controls adjusted by age, race and BMI. For PCB-156 and PCB-183 congeners, a significantly higher risk was found (OR=1.9, 95% CI 1.1-3.0) for the second exposure distribution tertile for PCB-156; and an OR=2.0 (95% CI 1.2-3.4) for the highest tertile of PCB-183.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

In another case-control study in Kingston and Toronto, Ontario, Canada, **[Aronson 2000]** non-cancerous breast fat tissue collected before treatment from 217 breast cancer incident cases and 213 biopsy controls was analyzed for 14 PCB congeners. A statistically significant association was found with breast cancer risk for PCB-105 (OR=3.17, 95% CI 1.51-6.68) and for PCB-118 (OR=2.31, 95% CI 1.11-4.78), with the fourth against the first quartile of the exposure distribution, and these effects have increased monotonically. Stronger associations were evident among premenopausal women for PCB-105 (OR=3.91, 95% CI 1.73-8.86) and for PCB-118 (OR=2.85, 95% CI 1.24-6.52) for the highest exposure category.

Also in the case-control study of the Long Island Hospital in New York [**Muscat 2003],** they measured the association between the total PCBs concentrations detected in adipose tissue and the risk of breast cancer in the participating women with a diagnosis of cancer. The highest tertile of the total PCB concentration was associated with an increased risk of relapse (RR=2.9, 95% CI 1.02-8.2) compared to the lower tertile.

Reviews and meta-analyses

In the reviews **[Brody 2007; Mouly 2016; Rodgers 2018]** the evidence for an association between PBCs and breast cancer risk in the general population is not conclusive.

But in a recent meta-analysis on exposure to different PCB congeners and breast cancer risk **[Zhang 2015]** the results of several epidemiological studies based on the classification of different congeners PCBs **[Wolff 1995]** in three different groups were evaluated; the PCBs belonging to group I with estrogenic effects, the PCBs belonging to group II with anti-estrogenic effects, the PCBs belonging to group III inducers of specific genes such as CYP1A and CYP2B on women with breast cancer. The results showed that the risk of breast cancer was statistically associated with the total exposure of the PCBs **[Zhang 2013]** to the PCBs congeners belonging to group II (with anti-estrogenic effects and immunotoxic congeners, similar to dioxin) and to groups of PCBs congeners belonging to group III (inducers Phenobarbital, CYP1A and CYP2B, biologically persistent), but not to PCBs congeners belonging to group I (potentially estrogenic).

However, in another meta-analysis **[Leng 2016]** the association between nine PCBs congeners and breast cancer was separately analyzed; the results of the meta-analysis imply that the PCB-99, PCB-183 and PCB-187 would increase the risk of breast cancer and have also found a significantly elevated risk in breast cancer among individuals with higher plasma/fat concentrations levels of PCB- 99, PCB-183 and PCB-187.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

References PCBs

- Gammon MD, Wolff MS, Neugut AI, et al. (2002) Environmental toxins and breast cancer on Long Island. II. Organochlorine compound levels in blood. Cancer Epidemiol Biomark Prev 11(8):686–697.
- Hoyer AP, Jørgensen T, Brock JW, Grandjean P (1998) Organochlorine exposure and risk of breast cancer. Lancet 352(9143):1816–1820.
- Hoyer AP, Jorgensen T, Grandjean P, Hartvig HB (2000). Repeated measurements of organochlorine exposure and breast cancer risk (Denmark). Cancer Causes Control, 11(2):177–84.
- Hoyer AP, Jørgensen T, Rank F, Grandjean P (2001) Organochlorine exposures influence on breast cancer risk and survival according to estrogen receptor status: a Danish cohort nested case-control study. BMC Cancer 1:8.
- Falck F Jr, Ricci A Jr, Wolff MS, Godbold J, Deckers P (1992). Pesticides and polychlorinated biphenyl residues in human breast lipids and their relation to breast cancer. Arch Environ Health, 47(2):143–6.
- Krieger N, Wolff MS, Hiatt RA, et al (1994). Breast cancer and serum organochlorines: a prospective study among white, black, and Asian women. J Natl Cancer Inst, 86(8):589–99.
- Demers A, Ayotte P, Brisson J, Dodin S, et al (2000). Risk and aggressiveness of breast cancer in relation to plasma organochlorine concentrations. Cancer Epidemiol Biomarkers Prev,9(2):161–6
- Demers A, Ayotte P, Brisson J, Dodin S, Robert J, Dewailly E (2002). Plasma concentrations of polychlorinated biphenyls and the risk of breast cancer: a congener- specific analysis. Am J Epidemiol,155(7):629–35.
- Wolff MS, Zeleniuch-Jacquotte A, Dubin N, Toniolo P (2000). Risk of Breast Cancer and Organochlorine Exposure. Cancer Epidemiology Biomarkers Prev, 9(3):271–7.
- Dorgan JF, Brock JW, Rothman N, Needham LL, Miller R, Stephenson HE Jr et al. (1999). Serum organochlorine pesticides and PCBs and breast cancer risk: results from a prospective analysis (USA). Cancer Causes Control, 10(1):1–11.
- Helzlsouer KJ, Alberg AJ, Huang HY, Hoffman SC, Strickland PT, Brock JW et al. (1999). Serum concentrations of organochlorine compounds and the subsequent development of breast cancer. Cancer. Epidemiol Biomarkers Prev, 8(6):525–32.
- Zheng T, Holford TR, Tessari J, Mayne ST, Owens PH, Ward B et al. (2000a). Breast cancer risk associated with congeners of polychlorinated biphenyls. Am J Epidemiol, 152(1):50–8.
- Zheng T, Holford TR, Mayne ST, Tessari J, Ward B, Carter D et al. (2000b). Risk of female breast cancer associated with serum polychlorinated biphenyls and 1,1-dichloro- 2,2′-bis(p-chlorophenyl)ethylene. Cancer Epidemiol Biomarkers Prev, 9(2):167–74.
- Holford TR, Zheng T, Mayne ST, Zahm SH, Tessari JD, Boyle P (2000). Joint effects of nine polychlorinated biphenyl (PCB) congeners on breast cancer risk. Int J Epidemiol, 29(6):975–82.
- Zhang Y, Wise JP, Holford TR, Xie H, Boyle P, Zahm SH et al. (2004). Serum polychlorinated biphenyls, cytochrome P-450 1A1 polymorphisms, and risk of breast cancer in Connecticut women. Am J Epidemiol, 160(12):1177–83.
- Rusiecki JA, Baccarelli A, Bollati V, Tarantini L, Moore LE, Bonefeld-Jorgensen EC (2008). Global DNA hypomethylation is associated with high serum-persistent organic pollutants in Greenlandic Inuit. Environ Health Perspect, 116(11):1547–52.
- Rubin CH, Lanier A, Kieszak S, Brock JW, Koller KR, Strosnider H et al. (2006). Breast cancer among Alaska Native women potentially exposed to environmental organochlorine chemicals. Int J Circumpolar Health, 65(1):18–27.
- Gatto NM, Longnecker MP, Press MF, Sullivan-Halley J, McKean-Cowdin R, Bernstein L (2007). Serum organochlorines and breast cancer: a case-control study among African-American women. Cancer Causes Control, 18(1):29–39.
- Itoh H, Iwasaki M, Hanaoka T, Kasuga Y, Yokoyama S, Onuma H et al. (2009). Serum organochlorines and breast cancer risk in Japanese women: a case-control study. Cancer Causes Control, 20(5):567-80.
- Ward EM, Schulte P, Grajewski B, Andersen A, Patterson DG Jr, Turner W et al. (2000). Serum organochlorine levels and breast cancer: a nested case-control study of Norwegian women. Cancer Epidemiol Biomarkers Prev, 9(12):1357–67.
- Laden F, Hankinson SE, Wolff MS, Colditz GA, Willett WC, Speizer FE et al. (2001a). Plasma organochlorine levels and the risk of breast cancer: an extended follow-up in the Nurses' Health Study. Int J Cancer, 91(4):568–74.
- Raaschou-Nielsen O, Pavuk M, Leblanc A, Dumas P, Philippe Weber J, Olsen A et al. (2005). Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. Cancer Epidemiol Biomarkers Prev, 14(1):67–74.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

- Liljegren G, Hardell L, Lindström G, Dahl P, Magnuson A (1998). Case-control study on breast cancer and adipose tissue concentrations of congener specific polychlorinated biphenyls, DDE and hexachlorobenzene. Eur J Cancer Prev, 7(2):135–40.
- Moysich KB, Ambrosone CB, Vena JE, Shields PG, Mendola P, Kostyniak P et al. (1998). Environmental organochlorine exposure and postmenopausal breast cancer risk. Cancer Epidemiol Biomarkers Prev, 7(3):181–8.
- Millikan R, DeVoto E, Duell EJ, Tse CK, Savitz DA, Beach J et al. (2000). Dichloro diphenyl dichloro ethene, polychlorinated biphenyls, and breast cancer among African-American and white women in North Carolina. Cancer Epidemiol Biomarkers Prev, 9(11):1233–40.
- Lucena RA, Allam MF, Costabeber IH , Villarejo MLJ , Navajas RFC. Breast cancer risk factors: PCB congeners.European Journal of Cancer Prevention 2001, 10, 117-119
- Lopez-Carrillo L, Blair A, López-Cervantes M, Cebrián M, Rueda C, Reyes R et al. (1997) Dichloro diphenyl trichloro ethane Serum Levels and Breast Cancer Risk: A Case Control Study from Mexico' Cancer Res, 57(17):3728–32.
- Holmes AK, et al. (2014) Case-control study of breast cancer and exposure to synthetic environmental chemicals among Alaska Native women. International Journal of Circumpolar Health 73, 257-60.
- Recio-Vega R, Velazco-Rodriguez V, Ocampo-Gómez G, Hernandez-Gonzalez S, Ruiz-Flores P, Lopez-Marquez F (2011). Serum levels of polychlorinated biphenyls in Mexican women and breast cancer risk. J Appl Toxicol, 31(3):270–8.
- Wolff MS, Toniolo PG (1995). Environmental organochlorine exposure as a potential etiologic factor in breast cancer. Environ Health Perspect, 103:Suppl 7: 141–5.
- Cohn BA, Terry MB, Plumb M, Cirillo PM (2012). Exposure to polychlorinated biphenyl (PCB) congeners measured shortly after giving birth and subsequent risk of maternal breast cancer before age 50. Breast Cancer Res Treat, 136(1):267–75.
- Zhang H, Liu L, Zhang P, Zhao Y, Wu X, Ni W. (2013). A case-control study on the relationship between Organochlorine and female breast cancer. Wei sheng yan jiu = Journal of hygiene research. 42(1):44–8.
- Charlier C, Pitance F, Plomteux G (2003). PCB residues in a breast cancer patient population. Bull Environ Contam Toxicol, 71(5):887–91.
- Charlier CJ, Albert AI, Zhang L, Dubois NG, Plomteux GJ (2004). Polychlorinated biphenyls contamination in women with breast cancer. Clin Chim Acta, 347(1-2):177–81.
- Stellman SD, Djordjevic MV, Britton JA, Muscat JE, Citron ML, Kemeny M et al. (2000). Breast cancer risk in relation to adipose concentrations of organochlorine pesticides and polychlorinated biphenyls in Long Island, New York. Cancer Epidemiol Biomarkers Prev, 9(11):1241–9.
- Aronson KJ, Miller AB, Woolcott CG, Sterns EE, McCready DR, Lickley LA et al. (2000). Breast adipose tissue concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk. Cancer Epidemiol Biomarkers Prev, 9(1):55–63.
- Muscat JE, Britton JA, Djordjevic MV, Citron ML, Kemeny M, Busch-Devereaux E et al. (2003). Adipose concentrations of organochlorine compounds and breast cancer recurrence in Long Island, New York. Cancer Epidemiol Biomarkers Prev, 12(12):1474–8.
- Wielsoe M, Kern P, Bonefeld-Jørgensen EC. (2017). Serum levels of environmental pollutants is a risk factor for breast cancer in Inuit: a case control study. Environmental Health1656.
- Arrebola JP, Belhassen H, Artacho-Cordón F, Ghali R, Ghorbel H, Boussen H, et al.(2015) Risk of female breast cancer and serumconcentrations of organochlorine pesticides and polychlorinated biphenyls: A case–control study in Tunisia Sci Total Environ 520:106–113.
- Leng L, et al. Polychlorinated biphenyls and breast cancer: a congener-specific meta-analysis. Environ Int. 2016;88:133–41.
- Zhang J, Huang Y, Wang X, Lin K, Wu K. (2015) Environmental Polychlorinated Biphenyl Exposure and Breast Cancer Risk: A Meta-Analysis of Observational Studies. PLoS One. 2015 Nov 10.
- Mouly TA, Toms LL. Breast cancer and persistent organic pollutants (excluding DDT): a systematic literature review. Environ Sci Pollut Res Int. 2016 Nov;23(22):22385-22407.
- IARC (2016). Polychlorinated biphenyls and polybrominated biphenyls. IARC Monogr Eval Carcinog Risk Chem Hum, Volume 107:1–503.
- Brody JG, Rudel RA, MS1 Michels KB, et al. (2007) Environmental Pollutants, Diet, Physical Activity, Body Size and Breast Cancer. CANCER Supplement June 15, 2007, Vol. 109, Number 12.
- Rodgers KM, Udesky JO, Rudel RA, Brody JG. Environmental chemicals and breast cancer: An updated review of epidemiological literature informed by biological mechanisms. Environmental Research 160, 152–182. 2018.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.2. Dichlorodiphenyltrichloroethane (DDT) and its metabolites (DDD, DDE) and others organochlorine compounds (Lindane, Hexachlorobenzene (HCB), Chlordane, ecc.)

Contaminants considered

Dichlorodiphenyltrichloroethane (DDT) is a synthetic industrial and household insecticide whose widespread use as a mosquito repellent and its long half-life made it a prominent environmental contaminant. Although DDT consequently helped lower incidence of malaria and typhoid, its use was banned in the US in 1972 for its effects on the environment and potential effects on human health. DDT and its main metabolites dichloro diphenyl dichloroethylene (DDE) and dichloro diphenyl dichloroethane (DDD) have endocrine disrupting properties.

Source

DDT is one of the most widely used insecticide in the world, successful in eradication of malaria from USA and other countries, though at an expense of devastating environmental problems and human health hazards. Between its first successful commercial applications in World War II and the severe usage restrictions in 1972, DDT was commonly used in households and agriculture because it does not wash off readily with water. It was the active ingredient in many aerosol fly sprays and a key ingredient in vegetable dusts and lawn and garden sprays.

Mean of exposure

Due to its extremely low solubility in water, DDT will be retained to a greater degree by soils and soil fractions with higher proportions of soil organic matter. It may accumulate in the top soil layer. Generally DDT is tightly sorbed by soil organic matter, but it (along with its metabolites) has been detected in many locations in soil and groundwater where it may be available to organisms.

Regulation and/or persistence

The EPA has assigned DDT, DDE, and DDD to Group B2, probable human carcinogens. The International Agency for Research on Cancer (IARC) has determined that DDT, DDE, and DDD are possibly carcinogenic to humans (Group 2B). The Department of Health and Human Services (DHHS) has determined that DDT, DDE, and DDD may reasonably be anticipated to be human carcinogens. In 1972, DDT was banned in the US, in 1978 in Italy. DDT is a very highly persistent organic pollutant in the environment that is readily adsorbed to soils and sediments.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Evidence on (breast) cancer risk

DDT and DDE have been suggested to be associated with an increased risk of female breast cancer because of their reported estrogenic activity and ability to induce p450 enzymes, which are intimately involved in steroid hormone metabolism. Different case-control studies adjusted for some or all of the standard reproductive and demographic risk factors for breast cancer, such as BMI, adult body-weight gain, family history, menopausal status, age at menarche, age at first birth, and lactation history have shown a correlation between blood DDT/DDE levels and development of breast malignancy.

IARC Monography

The International Agency for Research on Cancer (IARC) published a Monograph on the evaluation of carcinogenic risk to humans "DDT, Lindane and 2,4-D" Vol. 113 (2018). The IARC evaluation of the carcinogenicity and exposure of DDT, Lindane and 2,4-D and breast cancer was considered in many scientific articles.

The relationship between serum concentration of DDT and his main metabolite DDE, and breast cancer has been evaluated in different case-control studies.

For example, in the hospital-based case-control study conducted in 1994-1996 in Mexico **[Lopez-Carrillo 1997],** 141 cases aged between 20 and 79 years from participating hospitals and an equal number of controls without cancer were enrolled. Mean serum concentrations of 505.5 ng/g of lipids and 84.5 ng/g of lipids were reported for DDT (higher in controls) and DDE (higher in cases); the level of DDE found was not associated with the risk of breast cancer (OR=0.76, 95% CI 0.41-1.42); no risk data for DDT were reported.

The continuation of this study was carried out by **[Romieu 2000]** in Mexico by measuring serum DDE concentrations and DDT for a sub-sample of 120 breast cancer cases and 126 controls.

The concentration of DDE was higher in cases than in controls (but not statistically significant), whereas the concentration of DDT was not significantly higher in controls.

Another hospital-based case-control study was done in Rio de Janeiro, Brazil **[Mendonca 1999]**. In this study, women admitted with a diagnosis of breast cancer were enrolled as cases, and how hospital visitors are monitored without cancer. Cases and controls were interviewed in hospital and blood samples were obtained (before surgery for most cases) and analyzed for DDE and related compounds. Also in this case, no association between the risk of breast cancer and the increase in serum DDE was observed.

Also in another hospital study in New York, in the United States **[Wolff 2000a]** with cases of breast cancer and a control group of women undergoing surgery or biopsies for benign breast disease and a second control group of women without disease, concentrations of DDE and DDT determined in blood samples are not associated with the risk of breast cancer. In the same cohort, with an extended follow-up to 1994, only incident cases were considered.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Cases with at least three annual blood samples were included and serum DDE concentrations were adjusted for lipids; 110 cases and 123 controls were included in the analysis **[Wolff 2000b]**. also in this case the results did not confirm a significant increased risk of breast cancer in relation to DDE.
[Moysich 1998] studied the association between breast cancer in postmenopausal women with serum concentrations of different organochlorine compounds DDE, HCB, PCBs in a case-control study in the state of New York, USA. Data on serum DDE and risk factors for breast cancer were available for 154 cases and 192 controls, the risk of breast cancer was increased in the highest category of DDE exposure (OR=1.34, 95% CI 0.71-2.55), but there was no significant trend of response to exposure (p for trend = 0.25).

A well designed case control study is that of **[Krieger 1994]** performed in Northern California, United States, with a long follow-up period (1964-1990). Among the 2097 patients identified with breast cancer, 150 cases (50 whites, 50 blacks and 50 Asians) were randomly selected and matched with 150 controls per race, age, date of entry and date of follow-up. After adjustment for reproductive factors, menopausal status and body mass index (BMI), there was no significant association between breast cancer risk and DDE serum concentrations for all subjects, including ethnic, white subgroups, blacks and Asians.

Also others case control studies conducted in USA **[Hunter 1997; Dorgan 1999; Helzlsouer 1999; Millikan 2000; Zheng 2000; Laden 2001; Gammon 2002; Gatto 2007]**, in Québec, Canada **[Demers 2000]**, in Hanoi, Vietnam **[Schecter 1997]**, in Naples, Italy **[Dello Iacovo 1999]**, in Norway **[Ward 2000]**, in Denmark **[Hoyer 1998]**, in Belgium **[Charlier 2003]**, in Slovakia **[Pavuk 2003]**, in Egypt **[Soliman 2003]**, in Alaska **[Rubin 2006]**, in Japan **[Iwasaki 2008; Itoh 2009]** found no association between serum DDE and DDT levels and breast cancer risk.
In the Canary Islands, Spain, **[Boada 2012]** found that serum levels of lindane were not associated with breast cancer; this study was not matched by age, cases were significantly older than controls, and few women were exposed.

Different case-control studies were conducted on the relationship between the concentration of DDE or DDT in adipose tissue and the risk of breast cancer in the USA **[Zheng 1999]**, in Ontario, Canada **[Aronson 2000]**, in Spain **[Ibarluzea 2004]**, in Denmark **[Raaschou-Nielsen 2005]**, in Sweden **[Liljegren 1998]**: in all these studies there was no association between the concentration of DDE or DDT found in adipose tissue and risk of breast cancer.

Also **[van't Veer 1997]** in the EURAMIC (European community multicenter study) conducted in Germany, Netherlands, Northern Ireland, Switzerland, and Spain on breast cancer and exposure to DDE, after adjustment for BMI, age at first birth, current alcohol drinking, found no association between the concentration of DDE in adipose tissue and risk of breast cancer.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

There are also case-control and/or cohort studies within the IARC monograph that found a statistically significant association between DDT exposure and its metabolites such as DDE and breast cancer. For example, the cohort study of **[Wolff 1993]** on women's health in New York enrolled 14,290 women between 1985 and 1991. During this time, women who had been diagnosed with breast cancer 1-6 months after entering the study were defined as cases. The controls were randomly selected by all cohort members who were alive and cancer-free at the time of cancer diagnosis and the serum concentrations of DDE and DDT were determined in a total of 58 cases and 171 controls; the concentration of DDE were statistically higher in cases of breast cancer compared to control subjects (DDE in cases: 11.0 +/- 9.1 ng/ml; DDE in controls: 7.7 +/- 6, 8 ng/ml, p = 0.031).

In the Danish cohort study **[Hoyer 2000]** a statistically significant serum concentration of p,p'-DDT was associated with a significantly greater risk of triple breast cancer and the relation to a dose-response was evident (OR=3.6, 95% CI 1.1-12.2). Within the same cohort, a total of 161 cases with ER status information and 318 matched controls who were free of breast cancer were included in an analysis according to ER status **[Hoyer 2001]**. The observed increased breast cancer risk associated with exposure to dieldrin concentration (ng/ml) derived from women who developed an estrogen receptor negative (ERN): OR=7.6, 95% CI 1.3-46.1, I vs. IV quartile.

In the hospital-based case-control study **[Contreras 1998]** conducted in Bogota, Colombia, women with incident breast cancer and controls were recruited from 1995-1996. Blood samples for the cases were obtained before treatment; plasma DDE was higher in cases than controls. The risk of breast cancer is increased with the plasma concentration of DDE (OR=1.95, 95% CI 1.10-3.52) for the third compared to the first tertile.

[Charlier 2004] studied the association between breast cancer and serum DDT and concentration of DDE and hexachlorobenzene (HCB) among 231 cases recruited from a hospital surgery unit and 290 age-matched controls seeking cytological screening in Belgium in 2001-2002. DDT, DDE and HCB were measured in serum, and both concentration levels were low compared to other studies. The presence of both organochlorine compounds (DDE and HCB) in serum was statistically significant associated with an increased risk of developing breast cancer, with OR=2.21 (95% CI 1.41-3.48) for the DDE and OR=4.99 (95% CI 2.95-8.43) for HCB).

Several studies on human cancer on perinatal exposure to DDT have reported positive results.

In a prospective case–control study on young women (aged <20 years) in California, USA **[Cohn 2007]**, girls exposed to higher DDT/DDE levels were more likely to develop breast cancer than those with lower exposures, and risk increased with younger age at exposure. In this study they have found a positive trend between the serum level of p,p'-DDT (before it was banned, from 1959 to 1967) and the future risk of breast cancer (these women were mostly aged <20 years when DDT reached its peak of concentration). The number of identified cases (129) from the California registry, combined with the

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

same controls, had breast cancer before age 50. Significant positive associations were found with p,p'-DDT in blood samples before diagnosis (µg/l 8.09-13.90), OR=2.5 (95% CI 1.0-6.3).
In a subsequent study based on this cohort **[Cohn 2015]**, the incidence of breast cancer in 9300 daughters of women who provided blood samples in the original study was examined in relation to maternal prenatal exposure to DDT. Serum peritoneal serum concentration of o,p'-DDT was significantly associated with the risk of breast cancer in daughters in models adjusted for maternal lipids, overweight and breast cancer history.

Reviews and meta-analyses

In the reviews of **[Snedeker 2001; Calle 2002]** most of the nested case-control and case-control studies conducted since 1996 have not confirmed previous observations of a significant positive relationship between serum or adipose tissue levels of DDE or DDT and cancer risk at the otherwise. Most of these studies were conducted using Caucasian women from the United States, Canada or Europe. Because some studies have documented a tendency of black women in the United States to have higher serum or DDE levels than white women, further studies are needed to determine if black women have a higher risk of breast cancer associated with body loads of DDE. One of the reasons for the lack of association between DDE levels in blood or tissues and the risk of breast cancer may be the route of exposure to different forms of DDT and its various estrogenic metabolites.

The scientific reviews of **[Macon 2013; Gray 2017]** point out that there are also different epidemiological studies that have found positive associations between DDT or its metabolites such as DDE in serum and an increased risk in developing breast cancer, resume the prospective studies conducted by **[Cohn 2007 and 2015]** where girls exposed to higher levels of DDT / DDE were more likely to develop breast cancer than those with lower exposures and the risk increased with younger age following exposure.

In the two meta-analyzes on DDT and its metabolites DDE in general there are no big news and the risk of breast cancer **[Lopez-Cervantes 2004; Park 2014]** as they analyze the main scientific studies have already discussed in the previous paragraphs which are part of the most recent IARC monograph (2018) on the assessment of the human carcinogenic risk "DDT, Lindane and 2,4-D".
In the most recent meta-analysis **[Park 2014]** the scientific literature on exposure to DDT and its metabolites such as DDE and breast cancer risk was examined. In this meta-analysis, 35 case-control studies were analyzed and the overall result of the summary probability ratio (OR) for the identified studies was 1.03 (95% CI 0.95-12.12).

Subgroup analysis did not indicate a significant association between DDE exposure and breast cancer risk by design type, years of study, biological sample and geographical region of the study, except for case-population control studies based on estimation of DDE in serum published in the 90s.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

References DDT

- Lopez-Carrillo L, Blair A, López-Cervantes M, et al. (1997) Dichloro diphenyl trichloro ethane Serum Levels and Breast Cancer Risk: A Case Control Study from Mexico' Cancer Res, 57(17):3728–32.
- Contreras OP, Rodríguez-Villamil J, Posso-Valencia HJ, Cortez JE. (1998) Organochlorine exposure and breast cancer risk in Colombian women Cad Saude Publica, 14:Suppl 3:125–32.
- Moysich KB, Ambrosone CB, Vena JE, , et al. (1998) Environmental organochlorine exposure and postmenopausal breast cancer risk.Cancer Epidemiol Biomarkers Prev. 7(3):181–8.
- Mendonca GA, Eluf-Neto J, Andrada-Serpa MJ, Carmo PA, Barreto HH, Inomata ON et al. (1999). Organochlorines and breast cancer: a case-control study in Brazil. Int J Cancer, 83(5):596-600.
- Dello Iacovo R, Celentano E, Strollo AM, Iazzetta G, Capasso I, Randazzo G (1999). Organochlorines and breast cancer. A study on Neapolitan women. Adv Exp Med Biol, 472:57-66.
- Wolff MS, Zeleniuch-Jacquotte A, Dubin N, Toniolo P (2000a). Risk of breast cancer and organochlorine exposure. Cancer Epidemiol Biomarkers Prev, 9(3):271-7.
- Romieu I, Hernandez-Avila M, Lazcano-Ponce E, Weber JP, Dewailly E (2000). Breast cancer, lactation history, and serum organochlorines. Am J Epidemiol, 152(4):363-70.
- Schecter A, Toniolo P, Dai LC, Thuy LT, Wolff MS (1997). Blood levels of DDT and breast cancer risk among women living in the north of Vietnam. Arch Environ Contam Toxicol, 33(4):453-6.
- Pavuk M, Cerhan JR, Lynch CF, Kocan A, Petrik J, Chovancova J (2003). Case-control study of PCBs, other organochlorines and breast cancer in Eastern Slovakia. J Expo Anal Environ Epidemiol, 13(4):267-75.
- Soliman AS, Wang X, Di Giovanni J, Eissa S, Morad M, Vulimiri S et al. (2003). Serum organochlorine levels and history of lactation in Egypt. Environ Res, 92(2):110-7.
- Wolff MS, Toniolo PG, Lee EW, Rivera M, Dubin N (1993). Blood levels of organochlorine residues and risk of breast cancer. J Natl Cancer Inst, 85(8):648-52.
- Krieger N, Wolff MS, Hiatt RA, Rivera M, Vogelman J, Orentreich N (1994). Breast cancer and serum organochlorines: a prospective study among white, black, and Asian women. J Natl Cancer Inst, 86(8):589-99.
- Hunter DJ, Hankinson SE, a F, Colditz GA, Manson JE, Willett WC et al. (1997). Plasma organochlorine levels and the risk of breast cancer. N Engl J Med, 337(18):1253-8.
- Laden F, Hankinson SE, Wolff MS, Colditz GA, Willett WC, Speizer FE et al. (2001). Plasma organochlorine levels and the risk of breast cancer: an extended follow-up in the Nurses' Health Study. Int J Cancer, 91(4):568-74.
- Hoyer AP, Grandjean P, Jorgensen T, Brock JW, Hartvig HB (1998). Organochlorine exposure and risk of breast cancer. Lancet, 352(9143):1816-20.
- Hoyer AP, Jorgensen T, Grandjean P, Hartvig HB (2000). Repeated measurements of organochlorine exposure and breast cancer risk (Denmark). Cancer Causes Control, 11(2):177-84.
- Hoyer AP, Jorgensen T, Rank F, Grandjean P (2001). Organochlorine exposures influence on breast cancer risk and survival according to estrogen receptor status: a Danish cohort-nested case-control study. BMC Cancer, 1(1):8.
- Dorgan JF, Brock JW, Rothman N, Needham LL, Miller R, Stephenson HE Jr et al. (1999). Serum Organochlorine pesticides and PCBs and breast cancer risk: results from a prospective analysis (USA). Cancer Causes Control, 10(1):1-11.
- Helzlsouer KJ, Alberg AJ, Huang HY, Hoffman SC, Strickland PT, Brock JW et al. (1999). Serum concentrations of organochlorine compounds and the subsequent development of breast cancer. Cancer Epidemiol Biomarkers Prev, 8(6):525-32.
- Ward EM, Schulte P, Grajewski B, Andersen A, Patterson DG Jr, Turner W et al. (2000). Serum organochlorine levels and breast cancer: a nested case-control study of Norwegian women. Cancer Epidemiol Biomarkers Prev, 9(12):1357-67.
- Iwasaki M, Inoue M, Sasazuki S, Kurahashi N, Itoh H, Usuda M et al.; Japan Public Health Center-based Prospective Study Group (2008). Plasma organochlorine levels and subsequent risk of breast cancer among Japanese women: a nested case-control study. Sci Total Environ, 402(2-3):176-83.
- Raaschou-Nielsen O, Pavuk M, Leblanc A, Dumas P, Philippe Weber J, Olsen A et al. (2005). Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. Cancer Epidemiol Biomarkers Prev, 14(1):67-74.
- Cohn BA, Wolff MS, Cirillo PM, Sholtz RI (2007). DDT and breast cancer in young women: new data on the significance of age at exposure. Environ Health Perspect, 115(10):1406-14.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

- Cohn BA, La Merrill M, Krigbaum NY, Yeh G, Park JS, Zimmermann L et al. (2015). DDT exposure in utero and breast cancer. J Clin Endocrinol Metab, 100(8):2865-72.
- Wolff MS, Berkowitz GS, Brower S, Senie R, Bleiweiss IJ, Tartter P et al. (2000b). Organochlorine exposures and breast cancer risk in New York City women. Environ Res, 84(2):151-61.
- Boada LD, Zumbado M, Henriquez-Hernandez LA, Almeida-Gonzalez M, Alvarez-Leon EE, Serra-Majem L et al. (2012). Complex organochlorine pesticide mixtures as determinant factor for breast cancer risk: a population-based case-control study in the Canary Islands (Spain). Environ Health, 11(1):28
- Charlier C, Albert A, Herman P, Hamoir E, Gaspard U, Meurisse M et al. (2003). Breast cancer and serum organochlorine residues. Occup Environ Med, 60(5):348-51.
- Charlier C, Foidart JM, Pitance F, Herman P, Gaspard U, Meurisse M et al. (2004). Environmental dichlorodiphenyltrichlorethane or hexachlorobenzene exposure and breast cancer: is there a risk? Clin Chem Lab Med, 42(2):222-7.
- Zheng T, Holford TR, Mayne ST, Tessari J, Ward B, Carter D et al. (2000). Risk of female breast cancer associated with serum polychlorinated biphenyls and 1,1-dichloro-2,2′-bis(p-chlorophenyl)ethylene. Cancer Epidemiol Biomarkers Prev, 9(2):167-74.
- Itoh H, Iwasaki M, Hanaoka T, Kasuga Y, Yokoyama S, Onuma H et al. (2009). Serum organochlorines and breast cancer risk in Japanese women: a case-control study. Cancer Causes Control, 20(5):567-80.
- Gatto NM, Longnecker MP, Press MF, Sullivan-Halley J, McKean-Cowdin R, Bernstein L (2007). Serum organochlorines and breast cancer: a case-control study among African-American women. Cancer Causes Control, 18(1):29-39.
- Millikan R, DeVoto E, Duell EJ, Tse CK, Savitz DA, Beach J et al. (2000). Dichloro diphenyl dichloro ethene,polychlorinated biphenyls, and breast cancer among African-American and white women in North Carolina. Cancer Epidemiol Biomarkers Prev, 9(11):1233-40.
- Gammon MD, Wolff MS, Neugut AI, Eng SM, Teitelbaum SL, Britton JA et al. (2002). Environmental toxins and breast cancer on Long Island. II. Organochlorine compound levels in blood. Cancer Epidemiol Biomarkers Prev, 11(8):686-97.
- Demers A, Ayotte P, Brisson J, Dodin S, Robert J, Dewailly E (2000). Risk and aggressiveness of breast cancer in relation to plasma organochlorine concentrations. Cancer Epidemiol Biomarkers Prev, 9(2):161-6.
- van't Veer P, Lobbezoo IE, Martin-Moreno JM, Guallar E, Gomez-Aracena J, Kardinaal AF et al. (1997). DDT (dicophane) and postmenopausal breast cancer in Europe: case-control study. BMJ, 315(7100):81- 5.
- Liljegren G, Hardell L, Lindstrom G, Dahl P, Magnuson A (1998). Case-control study on breast cancer and adipose tissue concentrations of congener specific polychlorinated biphenyls, DDE and hexachlorobenzene. Eur J Cancer Prev, 7(2):135-40.
- Zheng T, Holford TR, Mayne ST, Owens PH, Ward B, Carter D et al. (1999). Beta-benzene hexachloride in breast adipose tissue and risk of breast carcinoma. Cancer, 85(10):2212–8.
- Aronson KJ, Miller AB, Woolcott CG, Sterns EE, McCready DR, Lickley LA et al. (2000). Breast adipose tissue concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk. Cancer Epidemiol Biomarkers Prev, 9(1):55–63.
- Ibarluzea JM, Fernandez MF, Santa-Marina L, Olea-Serrano MF, Rivas AM, Aurrekoetxea JJ et al. (2004). Breast cancer risk and the combined effect of environmental estrogens. Cancer Causes Control, 15(6):591-600.
- Rubin CH, Lanier A, Kieszak S, Brock JW, et al. (2006). Breast cancer among Alaska Native women potentially exposed to environmental organochlorine chemicals. Int J Circumpolar Health, 65(1):18-27.
- Lopez-Cervantes M, Sanchez LT, Tobías A, Carrillo LL. (2004) Dichlorodiphenyldichloroethane Burden and Breast Cancer Risk: A Meta-Analysis of the Epidemiologic Evidence. EHP, Vol. 112, N. 2.
- Park JH, Cha ES, Ko Y, Hwang MS , Hong JH, Lee WJ. (2014) Exposure to Dichlorodiphenyltrichloroethane and the Risk of Breast Cancer: A Systematic Review and Meta-analysis. Osong Public Health Res Perspect 5(2), 77 – 84.
- Macon MB, Fenton SE. (2013) Endocrine disruptors and the breast: early life effects and later life disease. J Mammary Gland Biol Neoplasia. ;18(1):43-61.
- Snedeker SM. (2001) Pesticides and Breast Cancer Risk: A Review of DDT, DDE, and Dieldrin EHP, Vol. 109, Supplement 1.
- Calle EE, Frumkin H, Henley SJ, Savitz DA, Thun MJ. (2002) Organochlorines and Breast Cancer Risk. CA Cancer J Clin. Sep-Oct;52(5):301-9.
- Gray JM , Rasanayagam S, Engel C and Rizzo J. (2017) State of the evidence 2017: an update on the connection between breast cancer and the environment. Environmental Health, 16:94.
- IARC (2018). DDT, LINDANE, AND 2,4-D. IARC Monogr Eval Carcinog Risk Chem Hum, Volume 113:1 503.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.3. Dioxin (TCDD)

Contaminants considered

TCDD is a chlorinated polycyclic hydrocarbon which is a product of combustion pollution and production chemicals, one of the most notorious organochlorines is 2,3,7,8-tetrachloridibenzo-p-dioxin (TCDD or dioxin). Agent Orange, the herbicide used by the United States Army in the Vietnam War, was found to be tainted by TCDD. TCDD is found in soil and water and, due to its lipophilic properties and the extended half-life of approximately 7-8 years in humans, TCDD bioaccumulates in animals and the environment. TCDD binds to the aromatic hydrocarbon receptor (AHR) and has anti-estrogenic properties. There are many other environmental contaminants that bind to AHR and their hazard levels are based on their effects related to TCDD or equivalent toxicity factor

Source

Dioxins are widespread environmental contaminants. They are produced by paper and pulp bleaching; incineration of municipal, toxic, and hospital wastes; certain electrical fires; and smelters (plants where metal is extracted from ores). They are also found as a contaminant in some insecticides, herbicides, and wood preservatives.

Mean of exposure

A particular dioxin that is likely to be carcinogenic to humans is called TCDD (2,3,7,8-tetrachlorodibenzo-pdioxin). The general population is exposed to low levels of TCDD primarily from eating dairy products, fish, and meat, including poultry. TCDD has a half-life of 7–9 years in humans. Due to the omnipresence of dioxins, all people have background exposure and a certain level of dioxins in the body, leading to the so-called body burden. Current normal background exposure is not expected to affect human health on average.

Regulation and/or persistence

TCDD was classified by IARC as a "known human carcinogen”. However, TCDD does not affect genetic material and there is a level of exposure below which cancer risk would be negligible. However, due to the high toxic potential of this class of compounds, efforts need to be undertaken to reduce current background exposure. In 2001, the Joint Food and Agriculture Organization of the United Nations (FAO)/WHO Expert Committee on Food Additives (JECFA) performed an updated comprehensive risk assessment of PCDDs, PCDFs, and “dioxin-like” PCBs. In order to assess long- or short-term risks to health due to these substances, total or average intake should be assessed over months, and the tolerable intake should be assessed over a period of at least 1 month. The experts established a provisional tolerable monthly intake (PTMI) of 70 picogram/kg per month. This level is the amount of dioxins that can be ingested over lifetime without detectable health effects.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Evidence on (breast) cancer risk

The few epidemiologic studies that examined the relationship between TCDD exposure and breast cancer risk are limited by small sample size and lack of individual exposure data. Breast cancer mortality and incidence increased in female workers employed in the production of TCDD-contaminated phenoxyherbicides.

Significantly increased mortality from breast cancer was reported in Chapaevsk (Samara Region, Russia) with a chemical plant known to be a source of TCDD because from 1967 to 1987 it produced hexachlorocyclohexane (lindane) and its derivatives **[Revich 2001]** so dioxins were detected in air, in soil, in the town's drinking water, in the cow's milk and in human blood samples. To assess cancer risk and reproductive health status, official medical statistical information was used. Chapaevsk women have a SMR higher risk overall due to breast cancer (SMR=2.1, 95% CI 1.6-2.7) and cervix cancer (SMR=1.8, 95% CI 1.0-3.1). Chapaevsk is an incredibly interesting site for further environmental-epidemiological research to assess the impact of dioxins on human health.

Two ecological studies **[Benko 2009]** analyzed the incidence of selected malignancies in two populations exposed to polychlorinated hydrocarbons, mostly TCDDs/Fs and PCBs, by comparing data available in the National Cancer Registry of the Slovak Republic and National Oncological Registry of the Czech Republic databases.
Neither TCDDs/Fs appear to contribute to the observed incidence of breast and prostate cancer in the Michalovce District and lower bladder cancer incidence in Uherske Hradiste District.

The purpose of a case-control study **[Viel 2008]** in the southwest border of Besançon (France) was to examine the association between dioxins emitted from a polluting urban solid waste incinerator (MSWI) and the risk of invasive breast carcinoma between women living in the area under the direct influence of the structure. They compared 434 incident cases of invasive breast cancer diagnosed between 1996 and 2002, and 2170 controls randomly selected from the 1999 population census. In this study, a dispersion model validated as a proxy for dioxin exposure was used, resulting in four exposure categories. Furthermore, by means of GIS (Geographic Information System) technology they used dioxin concentrations as relative data rather than absolute figures to estimate past exposure. These geographic exposure categories were evaluated by PCDD/F measurements from soil samples. No increased or decreased risk for breast cancer was found for any dioxin exposure category.

In the interesting ecological study on the population **[Verkasalo 2004]** carried out in the neighboring areas (<20.0 km) of the river Kymijoki (heavily contaminated) in the south of Finland, at the beginning of the eighties, the association between dibenzo-p- polychlorinated dioxins and dibenzofurans and different types of cancer, including breast cancer, particularly in the category of farmers living near the river, was investigated. They used the GIS method to map the registry data, in squares of 500 m × 500 m from

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

1981 to 2000. For breast cancer they found a RR=1.15 (95% CI 1.03-1.28) for those who live at a distance from the river less than 5 km (1.0-4.9 km).

In the ecological study done in Michigan, USA **[Dai 2008]** it was found that there is a strong association between high levels of incidence and aging of breast cancer near areas contaminated by high levels of dioxin in the soil. These results suggest that the increase in breast cancer incidence is spatially associated with dioxin contamination in the soil. Aging is a substantial factor in the development of breast cancer. Tests can be used for high surveillance and education, as well as formulating new study hypotheses for further research.

Another interesting ecological study on the population in the same area of Michigan, USA **[Guajardo 2009]** evaluated the previously determined geographical groups of incidence of breast and lung cancer among residents living near the Tittabawassee and Saginaw rivers using a new series of environmental factors. Breast cancer data was acquired by the Michigan Department of Health, along with data from point sources from the United States Environmental Protection Agency (EPA). The data sets were used to determine whether there is a spatial association between the risk of disease and the environmental contamination of several persistent contaminants, including dioxins. Data georeferencing techniques were used and statistical analyzes were performed to investigate the local risk of breast and lung cancer. The study shows that the neighborhoods (indicated with their zip code) in the immediate vicinity of the rivers were associated with a high risk of breast cancer, while an increased risk of lung cancer was detected among the neighborhoods in the immediate vicinity of pollution point sources and main motorways. In addition, the report on cancer statistics from the Institute of Public Health in Michigan shows that the counties of Midland and Saginaw (which contain the ZIP codes 48640, 48603, 48734, 48880, 48618 and 48732) had a higher rate of adjusted breast for the incidence of cancer in the period from 1994 to 2003 compared to the incidence rates in Bay County (located further away from the Tittabawassee river). High rates of breast cancer incidence were observed in ZIP code 48640 (OR=1.76, 95% CI 1.31–2.33), ZIP code 48603 (OR=1.65, 95% CI 1.23–2.20) and ZIP code 48734 (OR=1.88, 95% CI 1.34-2.63), all located near rivers and industrial facilities potentially responsible for environmental contamination of rivers. Finally, statistically significant cancer incidence clusters ($p \leq .001$) were observed among residents living near the rivers.
These findings are useful for researchers and government agencies for risk assessment, regulation and control of environmental contamination in floodplains.

IARC Monography
The International Agency for Research on Cancer (IARC), in the Monograph on the evaluation of carcinogenic risk to humans “Polychlorinated dibenzo-para-dioxins and Polychlorinated Dibenzofurans” Vol. 69 (1997) and in the Vol. 110F (2012), carried out an evaluation to investigate the association between TCDD exposure and cancer in humans.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

The IARC evaluation of the carcinogenicity of TCDD it is unique because more emphasis is given to all cancers combined, compared to specific site cancer, then results for all cancers do not refer to individual specific cancer sites. 2,3,7,8-Tetrachlorodibenzo-para-dioxin (TCDD) is the most potent dioxin. In IARC's evaluations, evidence of a causal association with TCDD exposure was considered strongest for lung cancer, non-Hodgkin's lymphoma (NHL), and soft-tissue sarcoma (STS) and all cancers combined in occupationally exposed populations **[IARC 2012]**. As a comparison, typical blood levels of TCDD in adults without known sources of exposure fall in the range of 1–10 ppt **[IARC 1997]**.

Reviews

Exposure to TCDD and breast cancer risk has been examined in detail by several review authors **[Laden 1998; Birnbaum 2003; Boffetta 2011; Jenkins 2012; Macon 2013]** without substantial evidence of increased risk. The review of **[Laden 1998]** on ecological studies highlights in particular the two studies following the accidental explosion at a chemical plant near Seveso, Italy, in 1976, where exposure to high levels of dioxin in the environment is assessed **[Bertazzi 1993 and 1997]** studying both mortality and incidence of cancer during the decade following the event and, in particular, no excess in incidence or mortality for breast cancer was observed in studies based on the population of Seveso residents. Though in a subsequent follow-up study **[Pesatori 2009]** of the incidence of cancer covering the period 1977-1996 including the subjects residing at the time of the accident in three contaminated areas with decreasing ground TCDD levels was updated (zone A, very high; zone B, high; zone R, low) also in an uncontaminated surrounding area of reference,the incidence of cancer did not differ from expectations in any of the contaminated areas, but for breast cancer an increased risk was found in zone A for females 15 years after the accident (five cases, RR=2.57, 95% CI 1.07-6.20).

The extension of the Seveso cancer incidence study confirmed an excess risk of neoplasia of the lymphatic and hematopoietic tissue in the most exposed areas. The high risk of breast cancer in females in Zone A after 15 years from the Seveso accident deserves a further in-depth investigation. In the review of **[Boffetta 2011]** two studies comparing TCDD level in breast adipose tissue among women with cancer and benign disease reported no statistically significant difference **[Hardell 1996; Reynolds 2005].** Overall, the evidence linking TCDD exposure to breast cancer risk is inconclusive in this review.

In the review of **[Macon 2013]** other population studies investigating the effects of high TCDD exposure in the Seveso area in Italy were examined: for example, the **[Warner 2002]** study found that women in the area have an increased risk of breast cancer even if not statistically significant, furthermore a subsequent follow-up assessment of that population **[Boffetta 2011]** found no association between TCDD exposure and breast cancer incidence or mortality.

Many of the girls who had been exposed to high levels of TCDD had not reached postmenopausal age (age at maximum risk of breast cancer in the normal population) at the time of data analysis, and it is possible that the reported results were underestimated. Therefore, a follow-up study to determine whether these women developed breast cancer would be useful to accurately assess the TCDD exposure-related health outcome.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

References Dioxin (TCDD)

- Warner M, Eskenazi B, Mocarelli P, et al. (2002) Serum dioxin concentrations and breast cancer risk in the Seveso Women's Health Study. Environ Heal Perspect.;110(7):625–8.
- Boffetta P, Mundt KA, Adami HO, Cole P, Mandel JS. TCDD and cancer: a critical review of epidemiologic studies. Crit Rev Toxicol. 2011;41(7):622–36.
- IARC (1997). Polychlorinated dibenzo-para-dioxins IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 69. Polychlorinated Dibenzo-para-dioxins and Polychlorinated Dibenzofurans. Lyon: IARC, 33–343.
- Laden F, Hunter DJ. (1998). Environmental risk factors and female breast cancer. Annu Rev Public Health 19:101–123.
- Bertazzi A, et al. (1993). Cancer incidence in a population accidentally exposed to 2,3,7,8- tetrachlorodibenzo-para-dioxin. Epidemiology 4:398–406.
- Bertazzi A, Guercilena S, Consonni D, Tironi A, et al (1997). Dioxin exposure and cancer risk: a 15-year mortality study after the "Seveso accident" Epidemiology.;8(6):646-52.
- Macon MB & Fenton SE. (2013) Endocrine Disruptors and the Breast: Early Life Effects and Later Life Disease. J Mammary Gland Biol Neoplasia 18:43–61.
- Hardell L, et al. (1996). Increased concentrations of octachlorodibenzo-p-dioxin in cases with breast cancer—Results from a case-control study. Eur J Cancer Prev 5:351–357.
- Reynolds P, et al. (2005). Regional variations in breast cancer incidence among California women, 1988–1997. Cancer Causes Control 16:139–150.
- Pesatori AC, Consonni D, Rubagotti M, Grillo P, Bertazzi PA. (2009) Cancer incidence in the population exposed to dioxin after the "Seveso accident": twenty years of follow-up. Environ Health. 15;8:39.
- Dai D, Oyana TJ. (2008) Spatial variations in the incidence of breast cancer and potential risks associated with soil dioxin contamination in Midland, Saginaw, and Bay Counties, Michigan, USA. Environmental Health, 7:49
- Revich B, Aksel E, Ushakova T, Ivanova I, Zhuchenko N, Klyuev N, Brodsky B, Sotskov Y. (2001) Dioxin exposure and public health in Chapaevsk, Russia. Chemosphere ;43(4-7):951-66.
- Bencko V, Rames J. et al. (2009) Human exposure to polyhalogenated hydrocarbons and incidence of selected malignancies – central European experience. Neoplasma 56, 4.
- Viel JF, Clément MC, Hägi M, Grandjean S, Challier B, Danzon A. (2008) Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure Int J Health Geogr. 28;7:4.
- Verkasalo PK, Kokki E, Pukkala E, Vartiainen T, Kiviranta H, Penttinen A, Pekkanen J. (2004) Cancer Risk Near a Polluted River in Finland. Environ Health Perspect. 112(9):1026–1031.
- Guajardo OA and Oyana TJ. (2009) A Critical Assessment of Geographic Clusters of Breast and Lung Cancer Incidences among Residents Living near the Tittabawassee and Saginaw Rivers, Michigan, USA Journal of Environmental and Public Health 1-16 pages.
- Jenkins S, Betancourt AM, Wang J, Lamartiniere CA. (2012) Endocrine-active chemicals in mammary cancer causation and prevention. J Steroid Biochem Mol Biol.;129(3-5):191-200.
- Birnbaum LS, Fenton SE.(2003) Cancer and developmental exposure to endocrine disruptors. EHP 111
- IARC (2012). Monographs on the Evaluation of Carcinogenic Risks to Humans, volume 100F: 2,3,7,8-Tetrachloro dibenzo para - dioxin, 2,3,4,7,8-pentachlorodibenzofuran, and 3,3′,4,4′,5-pentachloro biphenyl. Lyon, France: International Agency for Research on Cancer.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.4.Polycyclic Aromatic Hydrocarbons (PAH)

Contaminants considered

Polycyclic aromatic hydrocarbons (PAHs) are a group of several hundred organic chemically related compounds ubiquitous environmental pollutants generated primarily during the incomplete combustion of organic materials (e.g. coal, oil, petrol, and wood). Chemically the PAHs are comprised of two or more benzene rings bonded in linear, cluster, or angular arrangements. Although there are many PAHs, most regulations, analyses, and data reporting focus on only a limited number of PAHs, typically between 14 and 20 individual PAH compounds.

Source

Polycyclic aromatic hydrocarbons (PAHs) enter the environment through various routes, they are ecologically persistent, have different toxic effects based on their molecular structure on organisms through various actions. The general characteristics of PAHs are high melting and boiling points (therefore they are solid), low vapor pressure and very low aqueous solubility, also very soluble in organic solvents (highly lipophilic). PAHs also exhibit various functions such as light sensitivity, heat resistance, conductivity; It emits skills, resistance to corrosion and physiological action.

The atmosphere is the most important means of ubiquitous dispersal of PAHs in the environment, as PAHs are emitted into the atmosphere mainly by the incomplete combustion of organic matter. The combustion of sources can be natural (including volcanoes and forest fires) or anthropogenic (such as vehicle discharges, agricultural fires, power plants, coke plants, steel mills, foundries and other industrial sources). PAHs tend to be in higher concentrations in urban environments than in rural areas because most PAH sources are located in or near urban centers. Clearly the PAHs present in the atmosphere suffer the effect of dry and wet deposition that makes them deposit in the soil and in the water. Some of these PAHs come from nearby sources, such as the exhaust gases of adjacent cars, while others come from more distant sources and have been transported at various distances through the air.

Mean of exposure

The main route of exposure of PAHs in the population goes through the exposure of contaminants in the atmosphere both in the outdoor and indoor environment, in the soil passing to water through for example industrial effluents and accidental spills during the transport of oil at sea. In addition, PAHs can be contained by eating smoked or grilled foods (principally meat), smoking cigarettes or breathing smoke from open fireplaces with wood and/or pellets. **[ACGIH 2005]**. Routes of exposure include ingestion, inhalation and skin contact in occupational and non-occupational settings. Occupational exposure can also occur in workers who breathe exhaust gases, such as mechanics, street vendors, motor vehicle drivers, including workers in mining, metalworking or oil refining. People may be exposed to PAHs in the air and in the surface soil by direct inhalation, ingestion or skin contact.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Regulation and/or persistence

US government agencies have regulations that are relevant to workplace PAH and environmental exposures. In 2000, EPA **[U.S.EPA. 2000]** established environmental water quality criteria to protect human health from the carcinogenic effects of exposure to PAHs. The objective of these criteria was to set an undetectable level (zero concentration for carcinogenic PAHs in ambient water).

The EPA developed a maximum level of contaminants (MCL) for benzo (a) pyrene at 0.2 ppb. It is known that benzo (a) pyrene (BaP) is the most carcinogenic PAH.

The WHO orientation **[WHO 2003]** established the unit risk of lung carcinoma of BaP at 87x10-6 ng/m³ for lifetime exposure. According to many WHO member states, the reference values for BaP are between 0.1 and 1.3 ng/m³. According to the European Commission **[European Commission, Directive 2004/107/EC]**, the average annual target concentration should not be exceeded in the fraction of PM10 in 1 ng/m³ although unfortunately this objective has been exceeded in many European locations, in particular in the eastern countries.

Evidence on (breast) cancer risk

PAH have been shown to cause carcinogenic and mutagenic effects and are potent immunosuppressants. Effects have been documented on immune system development, humoral immunity and on host resistence. The most extensively studied PAHs are benzo(a)pyrene (BaP) and 7,12-dimethylbenzo anthracene (DMBA).

Like many other environmental chemicals that are associated with breast cancer risk, PAHs are lipophilic (fat-seeking) and are stored in the fat tissue of the breast. PAHs have been shown to increase risk for breast cancer through a variety of mechanisms.

DMBA (7,12-Dimethylbenz(a)anthracene) is commonly found in our environment and can be isolated from diesel exhaust, barbequed meat, tobacco smoke, overheated cooking oil, etc.

Several recent studies have reported that PAHs (estrogen-mimicking) can activate estrogen receptors (ER), either directly or indirectly, by producing estrogenic metabolites **[Pliskova 2005]**

Several population studies have examined the role of polycyclic aromatic hydrocarbons and the risk of breast cancer. The case-control study conducted at Long Island (LIBCSP) found a statistically significant increase for women with breast cancer with detectable levels of PAH-DNA adducts in the blood **[Gammon 2002, 2004]** and estimated an increase of breast cancer risk for women in the higher quintile than the lower one of PAH-DNA adducts.

In another study, elevated PAH-DNA adducts measured in breast cancer tissue were significantly associated with breast cancer (OR=2.56, 95% CI 1.05-6.24) **[Perera 2003]**.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Also the recent nested case–control study conducted in New York **[Shen 2017]** examined the association between polycyclic aromatic hydrocarbon (PAH)-albumin adducts in blood (80 cases and 156 controls). In this study women with detectable levels of PAH had a twofold association with breast cancer risk (OR=2.04, 95% CI 1.06–3.93) relative to women with non-detectable levels and there was a dose–response relationship with women with higher levels of PAH exposure having more than a fourfold increase in risk.

However, two cohort studies have not observed any association between PAHs and the risk of breast cancer that could be due to the measurement of non-carcinogenic PAH markers in urine where urinary measurements only reflect term exposure **[Lee 2010]** or in samples of blood **[Saieva 2011]**.

An increasing number of studies have found that exposure to PAHs may further increase the risk of breast cancer for women with high susceptibility to genetic variants involved in carcinoma metabolism, DNA repair and cell cycle control pathways **[Terry 2004]**. These stronger associations in subgroups defined by genetic variants suggest that even women with a higher risk of breast cancer based on family history would be at greater risk, but detecting interactions between environmental carcinogens and underlying risk requires a sufficient number of women. It has been hypothesized that women at greater risk of breast cancer from exposure to PAH are women who have a higher underlying inherent risk of breast cancer, predicted by their family history of cancer **[Terry 2004]**.

The genetic differences in DNA repair genes can lead to differences in cancer susceptibility and may change the impact of environmental exposures on cancer risk.
Soil measurements will likely reflect PAH concentrations in the air because the deposition of PAHs is proportional to airborne concentrations above the ground. The concentrations of BaP in the soil may be more stable and do not vary as much as the airborne exposures **[Shantakumar 2005]**.
The season of blood donation and smoking status are the strongest predictors of detectable adducts to PAH DNA. Other predictive factors are increased age, higher income, early age at menarche, fewer months of breastfeeding and BaP of the soil.
Halogenated and non-halogenated polycyclic aromatic hydrocarbons (HAH/PAH), such as polychlorinated dibenzo-p-dioxins and biphenyls and benzo (a) pyrene, have been recognized as significant and widespread environmental contaminants that induce AhR expression in target tissues, especially in the womb **[Pliskova 2005].**

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

IARC Monography

The International Agency for Research on Cancer (IARC) in the Monograph on the evaluation of carcinogenic risk to humans "Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures." Vol. 92 (2010) and in the Monograph "Bitumens and Bitumen Emissions, and Some N– and S-Heterocyclic Polycyclic Aromatic Hydrocarbons." Vol. 103 (2013) carried out an evaluation to investigate the association between PAH exposure and cancer in humans.

More than 60 individual PAHs were evaluated by IARC in Vol. 103 (2013) and others in Vol. 92 (2010). As noted in these Volumes, the mechanisms of action of the various PAHs vary: some are carcinogenic, some are not, and some are mutagenic, and some are not. There are at least three pathways by which these compounds are metabolized, and various PAHs are metabolized preferentially by various pathways. It is not clear how these compounds would be evaluated as a group because: (a) they have already been evaluated individually; and (b) the complex mixtures in which they occur are rarely analysed for the concentrations of more than a few PAHs. Polycyclic aromatic hydrocarbons are environmental contaminants that play an important carcinogenic role due to widespread population exposure; these persistent contaminants can be bio-transformed into reactive intermediates that form covalent PAH-DNA adducts that have mutagenic properties to initiate and/or promote tumorigenesis **[Phillips 1983; IARC 2010].**

Evidence suggests that exposure to PAHs has a causational effect on breast cancer in humans, yet this interaction is not clearly understood.

Review

The review of **[Korsh 2015]** examines the potential relationship between PAHs and breast cancer; it is also crucial to consider the geographic location and socioeconomic status of the patients. Studies conducted in Western New York in 2005 and 2007 have noted the necessity of investigating exposure to PAHs in relation to the location where patients resided during various critical periods in their lives, such as at the times of menarche and first birth **[Bonner 2005].**

Furthermore, in response to reports of high breast cancer mortality rates in the North eastern USA, a 1997 study of the region was conducted; this study found statistically significant clusters of breast cancer deaths in the New York City-Philadelphia metropolitan area, particularly in affluent suburban communities with ample access to health care **[Kulldorff 1997]**.

The researchers noted that they were unaware of any studies indicating greater exposure to PAHs in these communities, but it is nonetheless worth considering environmental conditions as a contributing factor to breast cancer mortality rates, as such factors can vary widely with location. Subsequent research conducted among breast cancer patients in one of such significant suburban cluster – Long Island, NY – reported that PAH-DNA adduct levels were higher, albeit not significantly, among this area's breast cancer patients than in the control population **[Gammon 2002]**.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

References Polycyclic Aromatic Hydrocarbons (PAH)

- Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Muti P, Trevisan M, Edge SB, Freudenheim JL: Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. Cancer Epidemiol Biomarkers Prev 2005; 14: 53–60.
- Kulldorff M, Feuer EJ, Miller BA, Freedman LS: Breast cancer clusters in the northeast United States: a geographic analysis. Am J Epidemiol 1997; 146: 161–170.
- IARC (2010). Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures, Monogr Eval Carcinog Risks Hum 92 (2010) 765–771.
- ACGIH (American Conference of Governmental Industrial Hygienists). Polycyclic aromatic hydrocarbons (PAHs) biologic exposure indices (BEI) Cincinnati. OH: American Conference of Governmental Industrial Hygienists; 2005.
- WHO (World Health Organization). Polynuclear aromatic hydrocarbons in drinking-water. Background document for development of WHO Guidelines for Drinking-water Quality; 2003 [WHO/SDE/WSH/03.04/59].
- European Commission, Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air, Off J L023 (2005) 0003–0016 [26/01/2005].
- U.S.EPA. 2000.Deposition of Air Pollutants to the Great Waters: Third Report to Congress. Office of Air Quality Planning and Standards. EPA-453/R-00-0005; June 2000.
- Phillips DH (1983) Fifty years of benzo(a)pyrene. Nature 303(5917): 468–472.
- Gammon MD, Santella RM, Neugut AI, Eng SM, Teitelbaum SL, Paykin A, Levin B, Terry MB, Young TL, Wang LW, Wang Q, Britton JA, Wolff MS, Stellman SD, Hatch M, Kabat GC, Senie R, Garbowski G, Maffeo C, Montalvan P, Berkowitz G, Kemeny M, Citron M, Schnabel F, Schuss A, Hajdu S, Vinceguerra V (2002) Environmental toxins and breast cancer on Long Island. I. Polycyclic aromatic hydrocarbon DNA adducts. Cancer Epidemiol Biomarkers Prev 11(8): 677–685.
- Gammon MD, Sagiv SK, Eng SM, Shantakumar S, Gaudet MM, Teitelbaum SL, Britton JA, Terry MB, Wang LW, Wang Q, Stellman SD, Beyea J, Hatch M, Kabat GC, Wolff MS, Levin B, Neugut AI, Santella RM (2004) Polycyclic aromatic hydrocarbon-DNA adducts and breast cancer: a pooled analysis. Arch Environ Health 59(12): 640–649.
- Lee KH, Shu XO, Gao YT, Ji BT, Yang G, Blair A, Rothman N, Zheng W, Chow WH, Kang D (2010) Breast cancer and urinary biomarkers of polycyclic aromatic hydrocarbon and oxidative stress in the Shanghai Women's Health Study. Cancer Epidemiol Biomarkers Prev 19(3):877–883.
- Saieva C, Peluso M, Masala G, Munnia A, Ceroti M, Piro S, Sera F, Bendinelli B, Pala V, Sieri S, Tumino R, Giurdanella MC, Panico S, Mattiello A, Vineis P, Polidoro S, Matullo G, Palli D (2011) Bulky DNA adducts and breast cancer risk in the prospective EPIC-Italy study. Breast Cancer Res Treat 129(2): 477–484.
- Terry MB, Gammon MD, Zhang FF, Eng SM, Sagiv SK, Paykin AB, Wang Q, Hayes S, Teitelbaum SL, Neugut AI, Santella RM (2004) Polymorphism in the DNA repair gene XPD, polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, and breast cancer risk. Cancer Epidemiol Biomarkers Prev 13(12): 2053–2058.
- Shantakumar S, Gammon MD, Eng SM, Sagiv SK, Gaudet MM, Teitelbaum SL, Britton JA, Terry MB, Paykin A, Young TL, Wang LW, Wang Q, Stellman SD, Beyea J, Hatch M, Camann D, Prokopczyk B, Kabat GC, Levin B, Neugut AI, Santella RM. Residential environmental exposures and other characteristics associated with detectable PAH-DNA adducts in peripheral mononuclear cells in a population-based sample of adult females. J Expo Anal Environ Epidemiol 2005;15:482–90.
- Perera F, Rundle A. Letter. Correspondence re: Gammon et al., Environmental Toxins and Breast Cancer on Long Island. I. Polycyclic Aromatic Hydrocarbon DNA Adducts. 11, 677–685, 2002 Cancer Epidemiol Biomarkers Prev. 12(1):75-6. 2003.
- Plíšková M, Vondráček J, Vojtěšek B, Kozubík A, Machala M. Deregulation of Cell Proliferation by Polycyclic Aromatic Hydrocarbons in Human Breast Carcinoma MCF-7 Cells Reflects Both Genotoxic and Nongenotoxic Events. Toxicological Sciences, Volume 83, Issue 2, 1 February 2005, Pages 246– 256.
- Shen J, Liao Y, Hopper JL, Goldberg M, Santella RM, Terry MB. Dependence of cancer risk from environmental exposures on underlying genetic susceptibility: an illustration with polycyclic aromatic hydrocarbons and breast cancer. Br J Cancer. 2017 Apr 25;116(9):1229-1233.
- Korsh J, Shen A, Aliano K, Davenport T. Polycyclic Aromatic Hydrocarbons and Breast Cancer: A Review of the Literature. Breast Care 2015;10:316‑318.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.5. Perfluoroalkyl substances (PFASs, PFOS, PFOSA and PFOA)

Contaminants considered

The perfluoroalkyl substances (PFASs) are a family of perfluorinated chemicals that consist of a carbon backbone typically 4–14 in length and a charged functional moiety (primarily carboxylate, sulfonate, or phosphonate). The two most widely known PFASs contain an eight-carbon backbone and include perfluorooctanoic acid (PFOA) and perfluorooctane sulfate (PFOS) that are persistent compounds enduring for prolonged periods in the environment following the release and therefore some companies have foreseen the interruption of production and the replacement of PFOA and PFOS, changing production processes, reducing the release and level of these compounds in their products.

PFOA is used as a polymerization aid in the manufacture of several types of fluoropolymers, which have been used in a wide variety of industrial and consumer products, such as Teflon and Gore-Tex. PFOA does not break down in most environments.

Source

The PFAA have been used since the '50s in various industrial and commercial applications thanks to their chemical-physical characteristics (resist high temperatures, greases and water) as consumer products for fabrics, carpets and clothing, paper linings for food use, non-stick cookware and fire resistant foams.

Mean of exposure

PFOA, along with PFOS and other PFASs, have been detected in a variety of environmental matrices worldwide. These include surface water, air, mud, soils, sediments and polar ice caps. For example, PFOA and PFOS were detected in the Tennessee River downstream of a fluorochemical production facility, in drinking water sources near production facilities in West Virginia.

PFOS and PFOA are persistent in the environment and are found in human blood, breast milk and liver with a half-life of 4-10 years. PFOA has the potential for environmental long-range transport, which makes emissions of PFOA a transboundary pollution problem.

PFOA is found at low levels in the serum of most people living in the United States, with higher levels seen in professionally exposed workers. Sources of exposure in the general population are not well established, but they probably include diet, drinking water, food packaging and household products.

Regulation and/or persistence

Besides PFOA also other substances in the PFASs group have properties of concern, which are targeted by the following international regulations: Perfluorinated carboxylic acids with a carbon chain of eleven to fourteen carbon atoms are also listed as substances of very high concern on the REACH

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

candidate list because of their very persistent and very bioaccumulative properties. Perfluorooctane sulfonic acid (PFOS) is listed as persistent organic pollutant (POP) in Annex B of the Stockholm Convention. The U.S. EPA (Environmental Protection Agency) has established health advisories for two chemical contaminants called PFOA and PFOS based on the agency's assessment of the latest peer-reviewed science. EPA's assessment indicates that drinking water with individual or combined concentrations of PFOA and PFOS below 70 parts per trillion (0.07 micrograms per liter µg/L) is not expected to result in adverse health effects over a lifetime of exposure.

Evidence on (breast) cancer risk

While several studies report on the association between breast cancer and PCBs, only a few have investigated associations with the PFAAs, but for their potential for environmental persistence, their long human half-life, and possible toxicity, in the scientific community there is a rising concern because PFAAs might be associated with human cancers and for this reason it is necessary to continue scientific research linked to exposure to these types of persistent contaminants.

The recent case-control study **[Wielsoe 2017]** on Inuit women from Greenland found significant, positive associations between breast cancer risk and PFAAs were observed.

The participants were asked to complete a questionnaire with information on reproductive history and lifestyle and to provide a blood sample. The serum levels of different persistent contaminants, among which 16 perfluoroalkyl acids (PFAAs), were determined.

Positive associations between breast cancer risk and PFOA was found in 2nd tertile of serum (ng/ml) levels (OR=1.26, 95% CI 1.01-1.58) and in 3rd tertile (OR=2.64, 95% CI 1.17-5.97).

The associations indicate that environmental exposure to PFOA can be a factor increasing the risk for breast cancer in Inuit women.

IARC Monography

The International Agency for Research on Cancer (IARC) in the Monograph on the evaluation of carcinogenic risk to humans "Some chemicals used as solvents and in polymer manufacture." Vol. 110 (2016) carried out an evaluation to investigate the association between PFOA exposure and cancer in humans. The results showed limited evidence of the carcinogenicity of PFOA and a positive association with cancers of the testes and kidney. IARC has classified PFOA as possibly carcinogenic to humans (Group 2B). Few population-based case-control studies were available that examined PFOA serum concentrations in relation to various types of cancer, in particular breast cancer.

Drinking-water is thought to have been the predominant source of intake of PFOA for a highly exposed population near a production facility in West Virginia, USA, studied by the C8 Science Panel **[Barry 2013; Vieira 2013]** where both surface water and groundwater were contaminated by water and air emissions from the facility.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

[Barry 2013] examined incident cancers occurring in 1992–2011 through the state cancer registries or medical record review. The total sample size was 32.254, of whom 3713 (11.5%) had worked at some time in the production plant. Individual-level data on residential history, drinking-water source, and tap-water consumption were obtained from the questionnaires administered in 2005–2011 in a cohort identified as a result of a lawsuit brought by residents of the area surrounding the fluoropolymer production plant in West Virginia. PFOA exposure was associated with kidney and testicular cancer in this population but for breast cancer there is no association between exposure to cumulative PFOA serum concentration (RR=0.93, 95% CI 0.88–0.99).

Using a case-control design **[Vieira 2013]** examined incident cancers occurring in 1996-2005, using West Virginia and Ohio state cancer registries. Cases living in 13 counties around the fluoropolymer production plant were identified; analyses were limited to 18 cancer types that were of a-priori interest, or that had at least 100 cases in each state. The controls for each analysis were all other cancer types, excluding cancers of the kidney, liver, pancreas, and testes. In one set of analyses, residence at time of diagnosis was used to assign study participants to specific water districts in Ohio and West Virginia. Breast cancer was no associated with PFOA serum concentrations (µg/l) in all different category (low to very high) compared with cases living in unexposed areas.

A case-control study in Greenland **[Bonefeld-Jorgensen 2011]** examined risk of cancer of the breast in relation to PFOS and PFSA exposure and found an association between breast cancer and PFOS (OR=1.03, 95% CI 1.001-1.07) and ΣPFSA (OR=1.03, 95% CI 1.00-1.05).

This case control study is very small (31 cases and 115 controls of incident cases of cancer of the breast in Greenland in 2002-2003) and for the working group of the IARC is uninformative because of the small sample size resulting from the high proportion of missing covariate data.

A later prospective Danish study **[Bonefeld-Jorgensen 2014]** evaluates the association between serum levels of PFAS in pregnant Danish women and the risk of premenopausal BC during a follow-up period of 10–15 years. An increased risk for the highest perfluorooctane sulfonamide (PFOSA) was found by a logistic regression analyses of the quintiles and the association with breast cancer showed a stronger significant correlation for PFOSA in the 5th quintile (RR=2.40, 95% CI 1.20-4.83) where the association is strongest among the young women below fourty years of age, but the overall results of this study suggest no strong or coherent association between breast cancer occurrence and the measured PFAS.

Another recent case-control study **[Ghisari 2014]** is based on environmental exposure to POPs including PFAS and breast cancer risk in inuit Greenlandic women and has studied the main effect of polymorphisms in genes involved in xenobiotic metabolism and in the biosynthesis of estrogens, CYP1A1, CYP1B1, COMT and CYP17, CYP19 and the BRCA1 founder mutation in relation to the risk of BC and to explore possible interactions between gene polymorphisms and serum PFASs levels on BC risk in greenland inuit women. The mutation and polymorphisms of BRCA1 founder in CYP1A1 (Val) and CYP17 (A1) may increase the risk of BC among Inuit women and the risk increases with higher serum PFOS and PFOA levels.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Reviews

The review of human PFOA exposure studies **[Steenland 2010]** is limited to assessing occupational exposure to US labor cohorts. In this review regarding the occupational exposure to PFOA for breast cancer no excess of breast cancer is recorded.

In addition, a more recent review **[Siddique 2016]** also takes up the case-control study on women coming from Greenland **[Bonefeld-Jorgensen 2011]** where, as we saw in the previous paragraph, the risk of breast cancer was found to be associated with the serum level both for PFOS and for the sum of PFSA, before and after adjustment for the confounding factors.

References Polycyclic PFAS

- Barry V, Winquist A, Steenland K (2013). Perfluorooctanoic acid (PFOA) exposures and incident cancers among adults living near a chemical plant. Environ Health Perspect, 121(11-12):1313‑8.
- Bonefeld-Jorgensen EC, Long M, Bossi R, Ayotte P, Asmund G, Kruger T, et al. (2011). Perfluorinated compounds are related to breast cancer risk in Greenlandic Inuit: a case control study. Environ Health, 10(1):88.
- Bonefeld-Jorgensen EC, et al. Breast cancer risk after exposure to perfluorinated compounds in Danish women: a case-control study nested in the Danish National Birth Cohort. Cancer Causes Control. 2014;25(11):1439‑48
- Vieira VM, Hoffman K, Shin HM, Weinberg JM, Webster TF, Fletcher T (2013b). Perfluorooctanoic acid exposure and cancer outcomes in a contaminated community: a geographic analysis. Environ Health Perspect, 121(3):318‑23. doi:10.1289/ehp.1205829 PMID:23308854
- Wielsoe M, Kern P and Bonefeld-Jørgensen CB. Serum levels of environmental pollutants is a risk factor for breast cancer in Inuit: a case control study. Environmental Health (2017) 16:56
- IARC (2016). Some chemicals used as solvents and in polymer manufacture. Volume 110. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Lyon: International Agency for Research on Cancer; 2016
- Ghisari M, Long M, Bonefeld-Jorgensen EC. Genetic polymorphisms in CYP1A1, CYP1B1 and COMT genes in Greenlandic Inuit and Europeans. Int J Circumpolar Health. 2013;72:211-13.
- Steenland K, Fletcher T, and Savitz DA. Epidemiologic Evidence on the Health Effects of Perfluorooctanoic Acid (PFOA). EHP, Vol. 118, N°8, 2010
- Siddique S, Kubwabo C, Harris SA. A review of the role of emerging environmental contaminants in the development of breast cancer in women. Emerging Contaminants 2 (2016) 204-219.
- https://ntp.niehs.nih.gov/ntp/ohat/pfoa_pfos/pfoa_pfosmonograph_508.pdf
- https://www.epa.gov/ground-water-and-drinking-water/drinking-water-health-advisories-pfoa-and-pfos

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.6. Triazine (atrazine, simazine, terbuthylazine and metabolites desethyl-atrazine, desethyl-terbuthylazine)

Contaminants considered

Triazine herbicides atrazine, simazine, terbuthylazine and desetyl-atrazine metabolites, desethyl-terbuthylazine are among the most widespread substances in the many years of investigations, both in surface and underground waters, with concentrations that often exceed the limit of 0.1 µg/l set for drinking water. Atrazine is used in agriculture as a selective pre and post-emergence herbicide for annual control of grass and broad-leaved weeds. It has been used on asparagus, bananas, citrus groves, coffee, conifers, forestry, orchards, prairies, herbaceous crops, guavas, macadamia orchards, corn (maize), oil palms, sorghum, sugar cane, pineapples, roses and screws. It has also been used as a soil sterilizer for airfields, parking lots and industrial sites and as an algaecide in swimming pools. Recently, many of the uses that contribute to water waste have been reduced or eliminated. In the European Union, where a limit of 0.1 µg/l has been set for all pesticide residues in drinking water and groundwater, the use of atrazine-containing herbicides has been limited mainly to agricultural uses on maize and on the sorghum.

Source

The continued use of atrazine as a primary herbicide for years has led to the contamination of aquifers and soil with it because atrazine is one of the most used herbicides with 76 million pounds applied every year as it is cheap and effective.

Mean of exposure

Atrazine is applied to agricultural fields or crops to kill weeds. It is also used near highways and railways for the same purposes. Some atrazines can enter the air after being applied to the ground. Some atrazines can also be washed from the ground due to rainfall and enter the surrounding areas, including streams, lakes or other watercourses. Some atrazines can migrate from the upper soil surface to deeper layers of the soil and enter the groundwater. People living near areas where atrazine has been applied to crops can be exposed through contaminated drinking water. Atrazine has been found in about 20 Superfund sites in the United States. People living near these sites may be exposed to higher levels of atrazine. If you are a worker working with atrazine, you may be exposed to higher amounts of atrazine. The government estimated that around thousand people could be exposed to atrazine in this way. Atrazine, one of the most widely used herbicides in the United States, is intentionally applied to crops, in particular corn, sugar cane, pineapple and sorghum. Therefore, people living near the areas where these crops are grown, especially agricultural workers and herbicide applicators applying atrazine, may be exposed to atrazine because it is used in agriculture.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

You may be exposed to atrazine if you are in the vicinity when the cultures are treated with atrazine, if you are involved in the application of atrazine in crops, or if you are close to other places where it is applied. Most often, atrazine is not found in high concentrations in the air, but can be found in higher concentrations in near-air facilities or in areas where it is applied to crops. You could also be exposed to atrazine by digging in the dirt that contains atrazine.

Children may be exposed to atrazine by playing in the dirt containing atrazine. You and your children may be exposed to atrazine if you drink water from wells contaminated with the herbicide. While it is used on many crops, it has not been found in many food samples and therefore only at very low levels. Therefore, it is very unlikely that it would be exposed to atrazine by eating foods.

Regulation and/or persistence

Atrazine was banned in the European Union in 2005 due to persistent contamination of groundwater caused by it and studies indicating its carcinogenic potential for mammary gland, prostate gland and also its correlation with ecological interruptions.

Other federal agencies that develop regulations for toxic triazine substances are the Environmental Protection Agency (EPA), the Occupational Safety and Health Administration (OSHA), the Food and Drug Administration (FDA), the Agency for Toxic Substances and the Register of diseases (ATSDR) and the National Institute for Occupational Safety and Health (NIOSH). Regulations and recommendations may be expressed in levels not to be exceeded in air, water, soil or food that are usually based on levels that affect animals; so they are regulated to help protect people.

Atrazine is currently under review for the re-registration of pesticides by the EPA. Therefore, the EPA can be contacted for more information on atrazine. OSHA has set a limit of 5 mg atrazine/m³ of laboratory air for an 8-hour working day. NIOSH recommends a standard for occupational exposure of 5 mg atrazine/m³ of laboratory air during a 10-hour shift to protect workers from a concern that atrazine may cause cancer. The EPA has set a maximum allowable amount of atrazine in drinking water of 3 µg/l. In addition, atrazine is designated as a limited-use pesticide, which means that only certified pesticide applicators can use atrazine.

Evidence on (breast) cancer risk

Atrazine is a known endocrine disruptor. It interferes with the pituitary-ovarian axis by decreasing the levels of prolactin and luteinization, changes that contribute to increase the cancer of the mammary gland. There is documented evidence of atrazine that causes dramatic damage to the reproductive structures of frogs, fish and other wildlife. Atrazine also induces an increased activity of the aromatase enzyme which causes an increase in estrogen levels that is directly linked to breast cancer. Epidemiological studies do not provide any support for a causal relationship between atrazine exposure and breast cancer, if triazine herbicides cause cancer in humans it is still a matter of debate.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

IARC Monography

The International Agency for Research on Cancer in the Monograph on the evaluation of carcinogenic risk to humans of Atrazine in the Vol. 73 (1999) downgraded atrazine from group 2B (possibly carcinogenic to humans) to group 3 (not classifiable with regard to its carcinogenicity to humans). The Working Group concluded that the animal mammary cancers associated with exposure to atrazine involve a non-DNA-reactive, hormonally mediated mechanism that is not relevant to humans.

Reviews

The purpose of the review **[Boffetta 2013]** was to critically review several epidemiologic population-based studies on atrazine exposure and the risk of cancer including breast cancer and to compare these studies with the conclusions of the Environmental Protection Agency and the scientific panel on the carcinogenicity of atrazine, to determine whether epidemiological studies support a causal relationship between atrazine and any specific cancer.

The EPA considered seven epidemiological studies including four ecological studies, two case-control studies and the potential agricultural health study (AHS).

According to the review of **[Boffetta 2013]**, four ecological studies have provided inconsistent and weak evidence of an association between atrazine and breast cancer **[Kettles 1997; Hopenhayn-Rich 2002; Muir 2004; Mills and Yang, 2006];** the only case-control study that found an association is that of **[Kettles 1997],** although the results in a better-designed follow-up study **[Hopenhayn-Rich 2002]** have not been replicated.

The Kentucky Environmental Study **[Kettles 1997]** evaluated triazine exposure and breast cancer incidence in Kentucky counties classified as having low, medium or high triazine exposure based on triazine concentration in groundwater from 1990 to 1991 and in surface waters from 1993 to 1994; for the hectares of wheat planted in 1970, 1980 and 1990; and for the total use of pesticides in 1979 in that area.

Breast cancer incidence rates for each county for the years 1991-1994 were obtained from the state registry. An OR of 1.14 (95% CI 1.08-19.10) and 1.20 (95% CI 1.13-1.28) was reported for mean and high levels of triazine exposure, respectively, for period 1993-1994.

This other important ecological study **[Hopenhayn-Rich 2002]** is an expansion of a previous investigation conducted in Kentucky **[Kettles 1997]** that have evaluated the association between atrazine exposure and breast cancer for the period 1993–1997 using the same methods.

On the basis of breast cancer incidence rates from the state registry for 1993–1997, there was no association between breast cancer and atrazine exposure. The OR were 1.01 (95% CI 0.96-1.05) and 0.98 (95% CI 0.93-1.02) for the highest and next to- highest atrazine exposure groups, respectively.

The ecological study in UK **[Muir 2004]** examined the distribution of breast cancer incidence and atrazine data in two agricultural counties with a high incidence of breast cancer (Lincolnshire and Leicestershire) to examine spatial clustering that might suggest a common exposure. Using a linear

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

regression analysis to examine the spatial association between incidence of breast cancer and atrazine, the use of atrazine was not significantly associated with breast cancer, nor there was any association between urban electoral departments in both provinces.

In this large study focused on an agricultural population **[Mills and Yang 2006],** where 23,513 breast cancer cases were analyzed, the authors' confidence in the results was reinforced by significant associations for established breast cancer risk factors.
The researchers evaluated the association between breast cancer rate data between the Latin Cancer Registry and the data on atrazine pounds applied by the California Pesticide Regulatory Department. For the highest group of atrazine use, the adjusted RR was 0.87 (95% CI 0.73-1.04) without dose-response relationship. No association between breast cancer incidence and atrazine data was found even in the ecological study of Latinas carrying out agricultural work in California.

Both the population-based case-control study in a high-use area of atrazine in Wisconsin and in North Carolina (AHS) **[McElroy 2007; Engel 2005]** found no association with breast cancer.

[McElroy 2007] used data from three case-control studies of chemical agricultural monitoring on breast cancer in rural Wisconsin, atrazine exposure values were assigned retrospectively to study participants. Exposure to atrazine was estimated using a weighted moving average on groundwater well data to calculate an interpolated value assigned to each individual; it emerges that there is no association between atrazine exposure and breast carcinoma (OR=1.1, 95% CI 0.9-1.5 for levels 1.0-2.9 ppb; OR=1.3, 95% CI 0.3-5.0 for levels above 3.0 ppb).

[Engel 2005] assessed the relationship between potential pesticide exposure and incidence of breast cancer among farmers, because the assessment of exposure to atrazine was less detailed in farmers' wives than in their husbands, so exposure data of husbands have been used as an indirect measure of wives' exposure. After adjustment for potential confounders, breast cancer was not associated with atrazine, based on the use of the wife (RR=0.7, 95% CI 0.4-1.2) or the use of the husband (RR=1.1, 95% CI 0.7-1.6), or for all cases of breast cancer or postmenopausal breast cancer (RR=0.4, 95% CI 0.1-1.0 for wife's exposure and RR=1.0, 95% CI 0.6-1.7 for husband's exposure).

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Also the cohort study of Iowa and North Carolina **[Beane Freeman 2011]** found no significant association between atrazine use and breast cancer according to nine exposed and 27 unexposed cases (RR=1.14, 95% CI 0.47-2.50) and no evidence of an increased risk for those above the median level of lifetime days of use.

Overall, the EPA considered both epidemiological and toxicological data to determine that "the database lacks evidence of an association between atrazine and breast cancer".

Also in the reviews of scientific literature exerted by **[Simpkins 2011; Sathiakumar 2011],** it can be seen that the conclusions are the same found in the most recent revision of **[Boffetta 2013]** and that there is no epidemiological evidence suggesting that exposure to atrazine is directly associated at the onset of breast cancer in women and to date there is still a lack of a plausible mode of action for the onset of breast cancer linked to exposure of atrazine in women.

References Triazine

- Kettles MK, Browning SR, Prince TS, Horstman SW (1997). Triazine herbicide exposure and breast cancer incidence: an ecologic study of Kentucky counties. Environ Health Perspect 105:1222–1227.
- McElroy JA, Gangnon RE, Newcomb PA, Kanarek MS, Anderson HA, Brook JV, et al. (2007). Risk of breast cancer for women living in rural areas from adult exposure to atrazine from well water in Wisconsin. J Expo Sci Environ Epidemiol 17:207–214.
- Beane Freeman L, Rusiecki JA, Hoppin JA, Lubin JH, Koutros S, Andreotti G, et al. (2011). Atrazine and cancer incidence among pesticide applicators in the Agricultural Health Study (1994–2007). Environ Health Persp 119:1253–1259.
- Engel LS, Hill DA, Hoppin JA, Lubin JH, Lynch CF, Pierce J, et al. (2005). Pesticide use and breast cancer risk among farmers' wives in the Agricultural Health Study. Am J Epidemiol 161:121–135.
- Hopenhayn-Rich C, Stump ML, Browning SR (2002). Regional assessment of atrazine exposure and incidence of breast and ovarian cancers in Kentucky. Arch Environ Contam Toxicol 42:127–136.
- Mills PK, Yang R (2006). Regression analysis of pesticide use and breast cancer incidence in California Latinas. J Environ Health 68:15–22.
- Muir K, Rattanamongkolgul S, Smallman-Raynor M, Thomas M, Downer S, Jenkinson C (2004). Breast cancer incidence and its possible spatial association with pesticide application in two counties of England. Public Health 118:513–520.
- Boffetta P, Adami HO, Berry C and Mandel JS. Atrazine and cancer: a review of the epidemiologic evidence European Journal of Cancer Prevention 2013, Vol 22 No 2
- Simpkins JW, Swenberg JA, Weiss N, Brusick D, Eldridge JC, Stevens JT, Handa RJ, Hovey RC, Plant TM, Pastoor TP, Breckenridge CB. Atrazine and Breast Cancer: A Framework Assessment of the Toxicological and Epidemiological Evidence Toxicol Sci. Oct;123(2):441-59. 2011.
- Sathiakumar, N., MacLennan, P. A., Mandel, J., and Delzell, E. A review of the epidemiologic studies of triazine herbicides and cancer. Crit. Reviews Toxicol. 41(Suppl. 1), 1–34. 2011.
- IARC (1999). Monographs on the evaluation of carcinogenic risks to humans. Evaluation and re- evaluation of some agents which target specific organs in rodent bioassays. 73: Lyon: IARC; pp. 59–113.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.7 – Cadmium (heavy metal)

Contaminants considered

Cadmium is a non-essential metal that is naturally present in the environment, even in the most pristine or remote ones. With current analytical facilities, it is well detectable in almost any environmental samples, e.g. rocks, soils, surface and rain water, plants and humans. The concentrations of Cd are typically in the parts per billion or parts per trillion (mass based parts) concentration range, i.e. trace levels. Older environmental concentration data always need to be treated with caution but currently there are numerous data to reliably identify Cd occurrence in the environment.

Source

Over 90% of the Cd in the surface environment comes from anthropogenic sources, including rock phosphate fertilizers, fossil fuel combustion ash, cement production waste, metallurgical works, municipal waste, sewage sludge and atmospheric deposition.

The main sources of non-occupational cadmium exposure in the general population include smoke, air, food and water contaminated with cadmium. Cadmium is also found naturally in natural phosphate, the main material used for mineral phosphate fertilizers (P). The long-term use of mineral fertilizers P has enriched the agricultural land with Cd.

Mean of exposure

Cadmium occurs in the environment as a divalent cation that can be absorbed by the biota. This means that this element can be transferred from soil and water to human food chains, with generally greater importance of terrestrial parts (food crops) than the aquatic environment (fish) in the human diet.

The evidence also indicates that the increase in total Cd concentrations of soil due to atmospheric deposition or use of phosphate fertilizers is associated with an almost proportional increase in Cd concentrations of crops, all other factors being constant. Cd smoke and diet are the main ways of exposure to Cd in pristine areas. Tobacco plants naturally contain high concentrations of Cd in the leaves, the amount of which varies considerably with the origin of the tobacco. Although cadmium is only one constituent of tobacco smoke, smoking does increase body cadmium burden **[Darbre 2006].** Food intake is the main source of Cd exposure for the general non-smoking population.

The levels of Cd in foods are typically high in offal, organs, equine products, crustaceans, cocoa, mushrooms and some green leafy vegetables. The incidence of these products in the average intake of Cd in the diet is low due to their low average consumption. In contrast, concentrations of Cd of cereals and potatoes have received considerable attention due to their greater impact on the intake of Cd by humans.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Both the natural and anthropogenic sources of cadmium, including industrial emissions and the application of fertilizers and sewage sludge to agricultural land, can lead to soil contamination and the increase of cadmium by crops and vegetables grown for the human consumption.

Cadmium transported from the air will also add the cadmium content in wild plants, berries, fungi and wild animals, which together add to oral human exposure. Although the emission of cadmium from point sources has decreased in recent years, there are still considerable quantities of cadmium in various products and constructions. This cadmium may eventually enter the human route of exposure.

The process of soil cadmium absorption by plants has improved at low pH, therefore, the decrease in soil pH due to environmental acidification, a recognized problem in Sweden, can further increase the cadmium content in the human exposure route.

When assessing human exposure to environmental cadmium, it is important to consider all routes of exposure, as well as factors affecting the level of exposure and dose, for example, lifestyle factors such as smoking, food consumption and nutritional factors which can influence the absorption of cadmium. Smoking is an important source of cadmium exposure.

Regulation and/or persistence

Classification and workplace exposure limits set by the American Conference of Governmental Industrial Hygienists for Cadmium are: A2 (substances suspected of being carcinogenic to humans), respirable fraction, TWA 0.002 mg/m³. While the exposure limit value established by the Occupational and Safety Health Administration (OSHA) is: 0.1 mg/m³.

Evidence on (breast) cancer risk

Recently, cadmium has been shown to be a chemical that destroys the endocrine system with estrogenic properties and a potential carcinogen of the prostate. In addition to being persistent and toxic, Cd is bioaccumulative with high concentrations occurring mainly in the kidney. The carcinogenic effects of cadmium on the prostate and breast in rats were probably due to the estrogenic properties of the heavy metal that allowed it to bind to the cellular estrogen receptors and thus imitate the actions of estrogen.

Two case–control studies **[McElroy 2006; Gallagher 2010]** have directly examined the association between urinary cadmium and breast cancer risk in US women by observing a significant trend in increased risk of breast cancer from elevated levels of urinary cadmium that also is independent of tobacco use; although smoking is a well-established source of cadmium exposure, the main route of cadmium exposure is food ingestion, in particular root vegetables. **[Gallagher 2010].**

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Another recent study in Lithuania **[Strumylaite 2011]** aimed to determine and compare the concentration of cadmium (Cd) in various biological media (breast tissue, urine and blood) of breast cancer and benign breast cancer patients. Women reported that the average cadmium levels in breast and urine tissue were significantly higher in breast cancer patients. Cancer patients with estrogen positive (ER) receptors had a significantly higher Cd concentration of breast tissue compared to patients with ER negative **[Strumylaite 2011].**

In Japan, exposure to environmental cadmium is relatively higher, even in non-polluted areas, than in other countries; in a recent case-control study in Japan **[Nagata 2013]** have found that a higher urinary cadmium level was associated with an increased risk of breast cancer among Japanese women. Women in the middle and highest tertiles of urinary cadmium levels showed significantly elevated OR of breast cancer relative to those in the lowest tertile; furthermore the trend of increased risk with increasing cadmium level was also statistically significant; urinary cadmium per 1.0 μg/g of creatinine increment was associated with an OR of 1.67 (95 % CI 1.39-2.01).

IARC Monography

The International Agency for Research on Cancer (IARC) in the Monograph on the evaluation of carcinogenic risk to humans in the Vol. 58 (1993) regarding possible causal associations between cadmium exposures and human cancers, reviewed and evaluated the available epidemiologic findings and other relevant information on cadmium exposures and concluded there was sufficient evidence in humans for the carcinogenicity of cadmium and cadmium compounds. So, cadmium has classified as a carcinogen for humans (group I) by the International Agency for Research on Cancer (IARC) on the basis of occupational studies of nickel-cadmium manufacturing, cadmium processing, cadmium-recovery plants, and copper-cadmium alloy plants. Most studies reported excess mortality from lung cancer among cadmium-exposed workers. There are very few case control and/or cohort studies correlating the exposure to cadmium and the risk of breast cancer, the most interesting case-control study for the size of the population is that of McElroy et al. 2006 described in detail in the previous paragraph.

Reviews

In the review of **[Jarup 1998]** diet is the main source of cadmium exposure in the general Swedish non-smokers population. The average daily intake is around 15 μg/day, but there are large individual variations due to differences in energy intake and eating habits**.**

Some recent studies indicate an association between cadmium exposure and human breast cancer, and it has been suggested that the effects of cadmium are mediated by the estrogen receptor independent of estradiol **[Morales 1994]**.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

A Finnish case–control study **[Antila 1996]** unexpectedly detected high concentrations of cadmium in breast tissue samples, which may indicate that cadmium-binding proteins exist in human breast tissue. The correlation of cadmium with estrogen receptors in breast cancer was suggestive. The authors cautioned that their results neither demonstrated nor denied the role of initiation, promotion or progression of cadmium in breast cancer. The cadmium concentrations found in the breast cancer patients did not differ statistically significantly from those of the healthy controls.

References Cadmium

- McElroy JA, Shafer MM, Trentham-Dietz A, Hampton JM, New comb PA (2006) Cadmium exposure and breast cancer risk. J Natl Cancer Inst 98:869–872.
- Gallagher CM, Chen JJ, Kovach JS (2010) Environmental cadmium and breast cancer risk. Aging 2:04– 814.
- Strumylaite L, Bogusevicius A, Abdrachmanovas O, Baranauskiene D, Kregzdyte R, Pranys D, Poskiene L (2011) Cadmium concentration in biological media of breast cancer patients. Breast Cancer Res Treat 125:511–517.
- Pan J, Plant JA, Voulvoulis N, Oates CJ, Ihlenfeld C. Cadmium levels in Europe: implications for human health. Environ Geochem Health. 2010 Feb;32(1):1-12.
- Morales PG, Saceda M, Kenney N, Kim N, Salomon DS, Gottardis MM, Solomon HB, Sholler PF, Jordan VC and Martin MB. Effect of cadmium on estrogen receptor levels and estrogen-induced responses in human breast cancer cells. J Biol Chem. 1994;269(24):16896-901.
- Nagata C, Nagao Y, Nakamura K, Wada K, Tamai Y, Tsuji M, Yamamoto S, Kashiki Y. Cadmium exposure and the risk of breast cancer in Japanese women. Breast Cancer Res Treat (2013) 138:235– 239.
- Jarup L, Berglund M, Elinder CG, Nordberg G, Vahter M. Health effects of cadmium exposure--a review of the literature and a risk estimate. Scand J Work Environ Health. 1998;24 Suppl 1:1-51.
- IARC (1993). Cadmium and cadmium compounds. In: Beryllium, cadmium, mercury and exposure in the glass manufacturing industry. Lyon: IARC, 1993: 119-237. IARC monographs on the evaluation of carcinogenic risks to humans, vol 58.
- World Health Organization (WHO). Cadmium. Geneva: WHO, 1992. Environmental health criteria, vol 134.
- World Health Organization (WHO). Cadmium. Geneva: WHO, 1992. Environmental health criteria, vol 135.
- Antila E , Mussalo-Rauhamaa H , Kantola M , Atroshi F , Westermarck T Association of cadmium with human breast cancer. The Science of the Total Environment [01 Jul 1996, 186(3):251-256]
- Darbre PD, Metalloestrogens: an emerging class of inorganic xenoestrogens with potential to add to the oestrogenic burden of the human breast. J. Appl. Toxicol. 2006; 26: 191–197

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

3.8. Trihalomethanes (THMs) (bromoform, bromo-dichloro-methane, dibromochloromethane, chloroform)

Contaminants considered

Trialomethane is a compound in which three hydrogen atoms of the methane molecule (CH4) are substituted with atoms of one or more halogens. If the three hydrogens are replaced with three atoms of the same halogen, then the compound is called haloform: chloroform (CHCl3), bromoform (CHBr3), iodoform (CHI3), fluoroform (CHF3).

Bromo-trihalomethanes are used as laboratory reagents, in the synthesis of organic compounds, as solvents. Bromoform has been used as a cough sedative. Chloroform is used in production of refrigerants and as a solvent; it was used in the past as an anesthetic. The main trihalomethanes are represented by chloroform, bromodichloromethane, dibromochloromethane, bromoform.

Source

Trihalomethanes (THM) are formed in drinking water mainly due to the chlorination of organic matter naturally present in raw water stocks but they are also found as pollutants introduced as such and then dispersed into the environment. The speed and the degree of formation of THM increase as a function of the concentration of chlorine and humic acid, temperature, pH and concentration of bromide ions.

Chloroform is the most common THM and the main DBP (disinfection by-products) in chlorinated drinking water. In the presence of bromides, the brominated THM are preferentially formed and the chloroform concentrations decrease proportionally. It is assumed that most of the THMs present in the water are eventually transferred into the air due to their volatility. For chloroform, for example, individuals may be exposed to high concentrations of chlorinated tap water during the shower. In Italy for example, the decree law D.L. 31/2001 imposes the limit of 30 µg / l for the sum of THM (more restrictive than the EU) for human consumption waters.

Mean of exposure

For volatile THMs, approximately equal contributions to total exposure come from four areas: drinking water ingestion; indoor air inhalation largely due to volatilisation from drinking water; inhalation and cutaneous exposure during showering or bathing; food ingestion; with everything except food exposure resulting mainly from drinking water. The exposure of indoor air to volatile THM is particularly important in countries with low rates of ventilation in homes and high rates of showering and bathing.

Four trihalomethanes (clloform, IARC 2B group), bromodichloromethane (IARC 2B group), dibromochloromethane (IARC group 3) and bromoform (IARC group 3), together with nine haloacetic acids based on bromine and chlorine, are the main by-products of chlorination on weight basis. The most recent studies quantified the exposure to trihalomethanes as a proxy for the whole mixture.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

Regulation and/or persistence

The most prevalent DBP in drinking water are trihalomethanes (THM), which are the only DBP group regulated in the EU with a maximum contaminant level of 100 µg/l. Several DBPs have been shown to be genotoxic in vitro assays and carcinogenic in animal experiments and the WHO International Agency for Research on Cancer (IARC) classifies chloroform and other widespread DBP as possible human carcinogens.

Evidence on (breast) cancer risk

Some studies indicate that chlorinated by-products in drinking water (e.g. chloroform) may contribute slightly to the risk of breast cancer, which is the basis for a possible correlation between some DBPs in drinking water and breast cancer. Among the few epidemiological studies on the exposure to DBP and BC, some have found a positive association **[Ribera 2018]**.

However, most of the studies concerning the effects of THMS on breast cancer show a negative correlation. For example**, [Marcus 1998]** conducted an ecological study that describes the association between the total levels of trihalomethane in water provided publicly and the incidence of female invasive breast cancer. Total trihalomethane levels were not materially associated with breast cancer risk, adapting to potential confounders. When stratified by race, the association observed for the aforementioned total trihalomethane category was not very different in black women than in white women. These ecological data are compatible with untreated or weak-treated trihalomethanes in drinking water related to breast cancer risk.

The retrospective cohort study **[Vinceti 2004]** conducted in northern Italy (Guastalla) is very interesting: they have studied the mortality of a cohort to which tap water was supplied with high chloroform and trihalomethane. In this study the risks of breast, ovarian and prostate cancer also tend to increase in subjects with higher socio-economic status, so the excess rates found in our exposed cohort cannot easily be attributed to life.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

IARC Monography

Association between the ingestion of chlorinated drinking water in excess with risk of cancer followed by mortality has been reported in several epidemiological studies in the IARC Monograph on the Evaluation of Carcinogenic Risk to Human "Chlorinated Drinking Water, Chloroform By-product, Some Other Halogenated Compound Cobalt and Cobalt Compound", Vol. 52 Lyon IARC, 1991 and in the "Monographs on the evaluation of some chemicals present in industrial and consumer products food and drinking-water" in the Vol. 101 (2013).

[Doyle 1997] conducted a study among women only and found a significantly increased risk and a dose-response relationship with levels of chloroform for all cancers included breast cancer.
In a cohort study in Finland **[Koivusalo 1997]** an excess risk of breast cancer (RR=1.33, 95% CI 1.02-1.74) was also observed in relation to surface water use; they have also identified a significantly increased risk among women for cancers of the urinary bladder, colon, oesophagus and breast with increasing mutagenicity of the water.

Reviews

In the review of **[Mohamadshafiee 2012]** some studies indicate that chlorination by-products in drinking water may slightly contribute to the risk of breast cancer; however, most of the studies concerning the effects of THMS on breast cancer show a negative correlation.

References Trihalomethanes

- Vinceti M, Fantuzzi G, Monici L, Cassinadri M, Predieri G, Aggazzotti G, 2004. A retrospective cohort study of trihalomethane exposure through drinking water and cancer mortality in northern Italy. Sci. Total Environ. 330, 47–53.
- Doyle TJ, Zheng W, Cerhan JR, Hong CP, Sellers TA, Kushi LH et al., 1997. The association of drinking water source and chlorination by-products with cancer incidence among postmenopausal women in Iowa: a prospective cohort study. Am. J. Public Health 87, 1168–1176.
- Koivusalo M, Pukkala E, Vartiainen T, Jaakkola JJ, Hakulinen T. 1997. Drinking water chlorination and cancer: a historical cohort study in Finland. Cancer Causes Control 8, 192–200.
- Marcus PM, Savitz DA, Millikan RC, Morgenstern H, 1998. Female breast cancer and trihalomethane levels in drinking water in North Carolina. Epidemiology 9, 156–160.
- Ribera FL, Gràcia-Lavedan E, Aragonés N, Pérez-Gómez B, Pollán M, Amiano P, Jiménez-Zabala A, Castaño-Vinyals G, Roca-Barceló A, Ardanaz E, Burgui R, Molina AJ, Fernández-Villa T, Gómez-Acebo I, Dierssen-Sotos T, Moreno V, Fernandez-Tardon G, Peiró R, Kogevinas M, Villanueva CM. Long-term exposure to trihalomethanes in drinking water and breast cancer in the Spanish multicase-control study on cancer (MCC-SPAIN). Environ Int. 2018 Mar;112:227-234. doi: 10.1016/j.envint.2017.12.031. Epub 2017 Dec 28.
- IARC (1991). Monograph on the Evaluation of Carcinogenic Risk to Human: Chlorinated Drinking Water, Chloroform By-product, Some Other Halogenated Compound Cobalt and Cobalt Compound, 52 Lyon IARC, 1991.
- IARC (2013) Monographs on the Evaluation of some chemicals present in industrial and consumer Products food and drinking-water, 101 Lyon IARC, 2013.
- Mohamadshafiee MR, Taghavi L. Health Effects of Trihalomethanes as Chlorinated Disinfection by Products: A Review Article. World Academy of Science, Engineering and Technology International Journal of Environmental and Ecological Engineering Vol:6, No:8, 2012

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

General conclusion

All scientific articles of population studies (case-control, cohort, etc.) we have researched and analyzed correlate the exposure to the main persistent contaminants (mainly POPs) and the risk of breast cancer; in our evaluation. we either found a statistically significant risk or no risk (denied studies) for breast cancer. To critically review the eligible articles, we evaluated inclusion criteria for participants, appropriateness of controls and reference groups, exposure assessment, whether the range of exposure (from no or low exposure to high) provides a strong comparison, control for confounding, the length of follow-up after exposure to allow for disease latency. We researched and analyzed **11 IARC monographs** on specific contaminants (e.g. PCBs, DDT, dioxins, atrazines, PAHs, PFOA, cadmium, trihalomethanes and its metabolites); we also reviewed **24 scientific reviews**, for a total of **130 scientific study articles** (see table 1), of these, 84 did not reveal a statistically significant risk for breast cancer (NO ASSOCIATION), while 46 scientific articles , found a statistically significant correlation between exposure to the contaminant and breast cancer (POSITIVE), even for a single molecule and/or congener (e.g. PCB).

Tab. 1 – Scientific Articles researched for different persistent pollutant family and breast cancer risk.

Contaminants	number of scientific articles	NO ASSOCIATION scientific articles	POSITIVE scientific articles
PCBs (209 congeners)	38	20	18
DDT, DDD, DDE and principal Organachlorines Compounds	40	33	7
Dioxins (TCDD)	13	8	5
PAHs	14	8	6
PFAA (PFOS, PFOSA and PFOA)	6	3	3
Triazine	8	7	1
Cadmium (Heavy Metal)	6	3	3
Trihalomethanes (THMs)	5	2	3
TOTAL Scientific Articles	130	84	46

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

We reviewed much of the literature on environmental pollutants and breast cancer and we found different scientific articles that supported positive associations with different congeners of polychlorinated biphenyls PCBs, DDE and PAHs in combination with certain genetic polymorphisms, and for exposure to cadmium, trialomethanes and dioxins, but most of the case-control studies examined do not find a statistically significant association between the exposure to contaminated persistent and the risk of breast cancer.

The persistent bioaccumulative organochlorines PCBs and DDT, DDD, DDE remained frequently studied, though nearly all studies still relied on biological measurements collected at or after diagnosis, which may capture exposures relevant to cancer promotion (but not earlier stages of breast cancer development). Two cohorts in particular provide unique opportunities to examine relevant periods of exposure. In the Child Health and Development Studies (CHDS), blood samples collected during pregnancy in the 1950-1960 allowed evaluation of breast cancer risks associated with DDT and PCBs exposures in early life, including in utero, during a period of high exposures **[Cohn 2007, 2012, 2015].**

Clearly, the correlation between exposure to environmental contaminants and breast cancer risk may also be associated with other persistent environmental contaminants. From the beginning of the study design, however, we focused on the main classes/families of persistent contaminants that are present in the environmental monitoring plans of the various Environmental Protection Agencies. In line with current European and national legislation in the field of water protection and of the soil these agencies carry out annual and various monitoring environmental campaigns on different environmental matrices (deep water and soil) across the European countries,.

Results are summarized in Annexes 1-8, in the form of tables, for the main classes of persistent contaminants.

Co-funded by
the Health Programme
of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 1 – 8 (Legend)

The results of the research of the 130 scientific articles are summarized in eight annexes in the form of tables for the different classes of persistent contaminants.

All the schematic tables have the same approach, starting from the name of the first three researchers of the scientific journal, what kind of design of the observational study (case-control, cohort, etc.) and the place, exposure to the main persistent contaminants, the years of interest of the study, the relative risk, the main comments relevant to the scientific article and the possible covariates and methods used.

The lines highlighted in green, are the scientific articles, which found a statistically significant risk of the three persistent contaminants considered in relation with breast cancer, we also took into account the articles where even one congener of persistent contaminant (eg PCB-203) and the significant relationship with the breast cancer.

Blanks are all articles where there is no statistically significant relationship between the persistent contaminant and the breast cancer.

Forward, we have included, in the ANNEX (1-8), the principal Monographs IARC for the classes of persistent pollutants analyzed, the main reviews and meta - analyzes always between the correlation of the persistent contaminants and the breast cancer.

Co-funded by the Health Programme of the European Union

EU funding disclaimer: This project has received funding from the 3rd European Union Health Programme 201-2020 under Grant Agreement PP-2-5-2016 (# 769767)

ANNEX 1 – References PCBs articles

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**Gammon MD**, Wolff MS, et al. Vol. 11, 686–697, **Cancer Epidemiology, Biomarkers & Prevention; 2002;**	Case-control study. Long Island, NY. **USA**	PCBs, DDE, Chlordane, Dieldrin.	Blood Samples	After diagnosis	1996-1997	No substantial elevation in breast cancer risk was observed in relation to the highest quintile of lipid-adjusted serum levels of PCBs	No statistically significant results for other contaminants measured.	Age, race, reproductive history, benign breast disease Interview response rates: cases, controls,
2	**Hoyer AP**, Jorgensen T, Rank F, et al. **BMC Cancer. 1:8. 2001.**	Nested case-control study. **Danish**	Total PCB; Dieldrin, HCB.	Blood Samples	Before diagnosis	1976 - 1978	Overall breast cancer survival in relation to serum concentrations of organochlorines by estrogen receptor status. Estrogen receptor positive ΣPCB: OR = 2.5 (1.1-5.7)	The results do not suggest that exposure to potential estrogenic organochlorines leads to development of an ERP breast cancer.	Breast cancer characteristics included tumor size, degree of spread, and stage of disease.
3	**Krieger N**, Wolff MS, Hiatt RA, et al. **Journal of the National Cancer Institute, Vol 86, No. 8, April 20. 1994.**	Nested case-control study; San Francisco Bay Area, **USA**	PCBs, DDE, DDT.	Serum	After diagnosis or at most 6 months before	1964 – 1990	Matched analyses found no differences in the case patients' and control subjects' serum levels of DDE or PCBs.	Conclusion: The data do not support the hypothesis that exposure to DDE and PCBs increases risk of breast cancer.	Race, age, date of entry, duration of follow up, BMI, age at menarche, menopausal status, ever pregnant No. and list of PCB congeners not provided.
4	**Falck FJ**, Ricci A Jr, Wolff MS, et al. **Arch Environ Health 47:143-146.1992.**	Case-control study; **Connecticut**	PCBs, DDT, DDE, HCB, HCH.	Adipose breast tissue	After diagnosis	1987-1992	The study by Frank Falck et al., showing higher chlorinated PCBs among 20 breast cancer cases, compared with 20 women with benign breast disease	There was approximately 1% increased risk for every 10 ppb of DDE and PCBs in adipose tissue.	The association persisted after controlling for age, smoking, and BMI.
5	**Rubin CH**, Lanier A, Kieszak S, et al. **International journal of circumpolar health** 65(1):18–27. **2006.**	Case-control study; **Alaska**	PCBs, DDT, DDE.	Serum Samples	Before diagnosis	1983-1987	OR = 1.43 (0.46, 4.47) for the highest tertile of DDE exposure and OR = 0.42 (0.07, 2.38) for the highest tertile of total PCB exposure.	Our results confirm exposure to organochlorines among Alaska. Native women but do not identify these exposures as a significant risk factor for breast cancer.	Parity, family history of breast cancer, race, triglycerides, cholesterol.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
6	**Recio-Vega R**, Velazco-Rodriguez V, Ocampo-Gómez G, et al. **J Appl Toxicol, 31(3):270-8. 2011.**	Case-control study. Comarca Lagunera, **Mexico**.	PCBs	Blood Samples	After diagnosis	2000-2007	They found 8 congeners PCBs (118, 128, 138, 170, 180, 195, 206 and 209) positively associated with BC risk.	statistically significant positive association between 8 congeners PCBs and BC risk	Age, age at menarche, lactation, menopausal status, BMI.
7	**Cohn BA**, Terry MB, Plumb M, et al. **Breast Cancer Res Treat 136(1):267–275. 2012.**	Prospective case-control study. Oakland, **California**	PCBs	Postpartum SERUM samples	Before diagnosis	1959 - 1998	Estimated net effects of PCB exposure on risk of breast cancer before 50 years of age. In contrast, PCB 203 was associated with a six fold increased risk (OR, 75th vs. 25th percentile = 6.3, 95 % CI 1.9, 21.7).	statistically significant positive association for PCB 203	Age, cholesterol, triglycerides, race, parity, lactation, BMI, year of blood collection.
8	**Arrebola JP**, Belhassen H, Artacho-Cordón F, et al. **Sci Total Environ 520: 106–113. 2015.**	Case–control study. **Tunisia**	PCB 138, PCB 153, PCB 180	Serum Samples	After diagnosis	2012	Concentrations of PCB 138, PCB 180, were significantly or borderline significantly higher in cases than in controls.	statistically significant positive association for PCB 138, 180.	Age, BMI, occupational class, residence, education, accumulated lactation time, parity, menopausal status, family history of breast cancer, total serum lipid
9	**Gatto NM**, Longnecker MP, Press MF, et al. Cancer Causes Control.18:29–39. **2007.**	Case–control study **USA**	PCBs, DDE, DDT.	Blood Samples	After diagnosis	1995-1998	PCBs: OR = 1.01, 95% CI = (0.63,1.63)	Breast cancer risk was not associated with increasing quintiles of lipid-adjusted PCBs or DDE	Age, BMI, lactation, lipid.
10	**Wolff MS**, Zeleniuch-Jacquotte A, et al. **Cancer Epidemiology Biomarkers Prev**, 9(3):271–7. **2000**	nested case-control study New York, **USA**	PCBs, DDE.	Serum Samples	6 months or more before breast cancer diagnosis	1994 - 1996	Quartiles of PCB concentration (ng/g lipid) 478–638 OR=1.55 (0.59–4.12) 639–876 OR=1.23 (0.49–5.08) > 876 OR=2.02 (0.76–5.37) P for trend = 0.23	Association not statistically significant	Age, menopausal status, date of blood collection (matching), age at menarche, number of pregnancies, age at first pregnancy, family history of breast cancer, lactation, height.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
11	**Itoh H**, Iwasaki M, Hanaoka T, et al. **Cancer Causes Control, 20(5):567–80. 2009.**	Case–control hospital study **Japan**	PCBs, DDT.	Serum Samples	After diagnosis	2001 - 2005	PCB-153 NR 0.40 (0.18–0.91) P for trend = 0.04 PCB-138 NR 0.61 (0.28–1.35) P for trend = 0.29 PCB-180 NR 0.29 (0.13–0.66) P for trend = 0.004	No association was identified between PCBs and risk of breast cancer	Age, total lipid concentration in serum, BMI, reproductive risk factors, medical history, area, lipid, menopause, smoking, fish and vegetable consumption, family history of breast cancer, age at first child birth, age at menarche, history of breast cancer screening, lactation.
12	**Zheng T**, Holford TR, Tessari J, et al. **Am J Epidemiol, 152(1):50–8. (2000a).**	Case–control study. Connecticut, **USA.**	PCBs	Serum Samples and adipose tissue	After diagnosis	1994–1997	PCBs (ppb) 396.0–562.9 79 exposed cases 0.6 (0.4–1.0) (ppb) ≥ 563.0 114 exposed cases 0.7 (0.4–1.1)	No association was identified between PCBs and risk of breast cancer	Age, BMI, structured interview; lipid-adjusted breast adipose tissue concentrations of 9 PCB congeners measured by GC (ng/g lipid)
13	**Zheng T**, Holford TR, Mayne ST, et al. **Cancer Epidemiol Biomarkers Prev, 9(2):167–74. (2000b).**	Case–control study. Connecticut, **USA.**	PCBs	Serum Samples and adipose tissue	After diagnosis	1995–1997	Total PCBs (ppb) 604.1–800.0 160 exposed cases 1.04 (0.76–1.45) (ppb) > 800.0 160 exposed cases 0.95 (0.68–1.32) P for trend = 0.41	NEGATIVE No association was identified between PCBs and risk of breast cancer	Age, BMI, structured interview; lipid- adjusted serum concentrations of 9 PCB congenerers measured by GC (ng/g lipid)
14	**Demers A,** Ayotte P, Brisson J, et al. **Cancer Epidemiol Biomarkers Prev, 9(2):161–6. 2000.**	Case–control study. Quebec City, **Canada.**	PCBs, DDE.	Serum Samples	After diagnosis	1994–1997	PCB-153 (µg/kg, lipid basis) 36.3–46.6, Hospital OR (95% CI)1.02 (0.54–1.94) Population OR (95% CI) 1.12 (0.66–1.88) 46.6–57.1, Hospital OR (95% CI) 0.99 (0.50–1.93) Population OR (95% CI) 0.94 (0.55–1.62)	Women diagnosed with breast cancer who had higher plasma concentrations of p,p'-DDE, b-HCH, oxychlordane, or trans-nonachlor were more likely to show both a large tumor (diameter > 2 cm) and axillary-lymph-node involvement.	Age, BMI, alcohol consumption, age at first cigarette, number of fertile years, age at first child, total breast feeding duration, use of oral contraceptive, use of hormone therapy, first-degree family history of breast cancer, history of benign breast disease, and time separating blood sampling from surgery.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
15	**Demers A**, Ayotte P, Brisson J, et al. **Am J Epidemiol, 155(7):629–35. 2002.**	Case–control study. Quebec City. **Canada.**	PCBs	Serum Samples	After diagnosis	1994–1997	PCB 105, 118, and 156 expressed as 2,3,7,8-tetrachlorodibenzo-p-dioxin toxic equivalents OR = 2.02, 95% CI: 1.24, 3.28; fourth vs. first quartile	There is a statistically significant association for PCB 105, 118, and 156 expressed as 2,3,7,8-tetrachloro dibenzo-p-dioxin toxic	Information on age, BMI, lifestyle, dietary habits, and reproductive history was obtained by telephone interview.
16	**Holford TR**, Zheng T, Mayne ST, et al. **Int J Epidemiol. Dec; 29(6):975-82. 2000.**	Case–control study. Connecticut. **USA.**	nine PCB congeners (74, 118, 138, 153, 156, 170, 180, 183, 187)	Breast adipose tissue samples	After diagnosis	1994–1997	Linear logistic model 10-ppb change in exposure PCB-183 OR=1.82 (95% CI; 1.12–2.98)	PCB-183 were associated with increased risk of cancer of the breast.	Age, BMI, reproductive risk factors, dietary fat intake, income, fat concentrations of PCB
17	**Zhang Y**, Wise JP, Holford TR, et al. **Am J Epidemiol, 160(12):1177–83. 2004.**	Case–control study. Connecticut. **USA**	PCBs	Serum Samples	After diagnosis	1999–2002	No significant association for CYP1A1 m1 or m4 genotype or in premenopausal women.	The CYP1A1 m2 genetic polymorphism was associated with increased risk of female breast cancer and may modify the relation between PCB exposure and breast cancer risk.	Age, BMI, lipid- adjusted serum concentrations of 9 PCB congeners. Genotyping of CYP1A1 m1, m2, and m4 by PCR-RFLP
18	**Rusiecki JA**, Holford TR, Zahm SH, et al. **Eur J Epidemiol, 19(8):793–801. 2004.**	Case–control study. Connecticut. **USA**	PCBs	Blood serum and Adipose tissue samples	After diagnosis	1994–1997	Tumours were apparent with concentrations of PCB-183 (third tertile vs first: OR, 2.4; 95% CI, 1.0–6.0, P for trend = 0.03, but data not otherwise shown).	Analyses for individual congeners did not show any association	Age, reproductive risk factors, BMI, family history of breast cancer in a first-degree relative.
19	**Laden F**, Hankinson SE, Wolff MS, et al. **Int J Cancer.** 91(4):568–74. **2001.**	nested case-control study **USA**	PCB and DDE	Plasma	After and before diagnosis	1976–1994	OR for breast cancer associated with PCBs was 0.94 (95% CI 0.73 to 1.21), and the equivalent OR for breast cancer associated with DDE was 0.99 (95% CI 0.77 to 1.27).	no significant associations were found between organochlorine levels in serum and breast cancer risk.	Age, family history of breast cancer, history of BBD, age at menarche, BMI, lipid, number of children, age at birth of first child, lactation.
20	**Ward EM**, Schulte P, Grajewski B, et al. **Cancer Epidemiol Biomarkers Prev. 9(12):1357-67. 2000.**	(nested case-control study **Norway**	PCBs	Serum	before diagnosis	1973 - 1991	The current study did not find any evidence for an association between organochlorine levels in serum and breast cancer.	no significant associations were found between organochlorine levels in serum and breast cancer risk.	Age, BMI, occupational category, lipid concentration in serum, year blood collected, age at blood collection. Age at first birth

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
21	**Helzlsouer KJ,** Alberg AJ, Huang H-Y, et al. **Cancer Epidemiol Biomarkers Prev. 8(6):525–32. 1999.**	nested case-control study **USA**	PCBs, DDE, DDT.	Serum	before diagnosis	1974 - 1994	The risk of developing breast cancer also tended to decrease with increasing levels of PCBs.	no significant associations were found between organochlorine levels in serum and breast cancer risk.	Adjustment for family history of breast cancer, BMI, age at menarche or first birth, and months of lactation, lipid concentration in serum.
22	**Dorgan JF,** Brock JW, Rothman N, et al. Cancer Causes Control.10(1):1–11. **1999.**	nested case-control study **USA**	PCBs and DDT.	Blood Sample	before diagnosis	1977 - 1987	Women with higher serum levels of other organochlorine pesticides and PCBs showed no increased risk of breast cancer.	Results of this study do not support a role for organochlorine pesticides and PCBs in breast cancer etiology.	Age, BMI, menopausal status, family history, estrogen use, parity, age at menarche
23	**Hoyer AP,** Grandjean P, Jørgensen T, et al.. **The Lancet. 352(9143):1816–20.1998.**	nested case-control study. **Denmark**	PCBs, Dieldrin, DDT.	Serum	After diagnosis	1976 – 1995	PCB OR= 1.11; (95% CI 0.70–1.77) Association not statistically significant with breast cancer	No association statistically significant was identified between PCBs, Dieldrin, DDT and risk of breast cancer	Age, weight, height, number of children, alcohol, smoking, lipid, physical activity, menopausal status, household income, marital status, education.
24	**Raaschou-Nielsen O,** Pavuk M, LeBlanc A, et al. **Cancer Epidemiol Biomarkers Prev. 14(1):67–74. 2005.**	nested case-control study **Denmark**	PCBs and others organochlorine compound.	Adipose breast tissue	Tissue was sampled up to 6.5 years before the breast cancer diagnosis	1993 – 1997	The results showed no higher risk of breast cancer among women with higher levels of any pesticides or polychlorinated biphenyls;	No association statistically significant was identified between PCBs and others organochlorine compound and risk of breast cancer	Age, education, BMI, lipid, alcohol, number of children, age at birth of first child, lactation, HRT, history of BBD.
25	**Zhang H**, Liu L, Zhang P, et al. **Journal of hygiene research. 42(1):44–8. 2013.**	case-control study **China**	PCB 28 and PCB 52, HCH, DDE.	Serum	/	2010 – 2011	PCBs OR= 7.46; 95% CI 2.44– 22.81	Organochlorines resides, including DDT, HCH and PCB, may increase women's risk of getting breast cancer.	Age, family history of breast cancer, history of BBD, age at menarche, lactation, menstrual cycle, time between menarche and prim parity, bean products intake.
26	**Charlier C,** Pitance F, Plomteux G. **Bull Environ Contam Toxicol, 71(5):887–91. 2003.**	case-control study **Belgium**	PCBs	Serum	After diagnosis	NR	PCB-101 and PCB-153 were significantly higher for cases than controls.	Further results are necessary	Age, age at menarche, menopause, HRT, parity, lactation, family history of breast cancer.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
27	**Charlier CJ,** Albert AI, Zhang L et al. **Clin Chim Acta. 347(1):177–81. 2004.**	case-control study **Belgium**	PCBs	Serum	After diagnosis	NR	PCB153 (ppb) OR = 1.8 (95% CI1.4– 2.5) < 0.0001	These results suggest that environmental exposure to PCBs may contribute to multifactorial pathogenesis of breast cancer.	Age, age at menarche, menopause, HRT, parity, lactation, family history of breast cancer.
28	**Lopez-Carrillo L,** Lopez-Cervantes M, Torres-Sanchez L et al. **Eur J Cancer Prev. 11(2):129–35. 2002.**	case-control study **Mexico**	PCBs , HCH, HCB.	Serum	Before and after diagnosis	1994 - 1996	OR 1.31; 95% CI 0.33, 5.21	This study lends no support to the case for a role for b-HCH, HCB or PCBs in breast cancer aetiology.	Age at menarche, number of children, age at first birth, lactation, lipid, family history of breast cancer, menopausal status, quetelet index.
29	**Moysich KB,** Ambrosone CB, Vena JE, et al. **Cancer Epidemiol Biomarkers Prev. 7(3):181–8. 1998.**	case-control study **USA**	PCBs, DDE, HCB.	Serum	After diagnosis	1986 - 1991	PCBs OR=1.66; 95% CI,1.07–2.88 for the combined second and third tertiles); association with total PCBs OR= 2.87; 95% CI, 1.01–7.29) and moderately chlorinated PCBs OR=3.57; 95% CI,1.10–8.60.	These results suggest that an increase in risk of postmenopausal breast cancer associated with environmental exposure to PCBs	Age, education, family history of breast cancer, parity, quetelet index, lactation, age at first birth, years since last pregnancy, fruit and vegetable intake, lipid.
30	**Millikan R,** DeVoto E, Duell EJ et al. **Cancer Epidemiol Biomarkers Prev. 9(11):1233–40. 2000.**	case-control study **USA**	PCBs, DDE	Plasma	After diagnosis	1993 – 1996	PCB exposure among African-American women (third tertile) OR=1.74; 95% CI,1.00–3.01) African-Americans with BMI > 34.2 (third tertile total PCBs, OR= 4.92; 95% CI,1.63–14.83	We observed no overall association between plasma levels of DDE and total PCBs but for breast cancer among African-American there is an association	Age, age-squared, race (all participants), lipid, menopausal status, BMI, parity / lactation, HRT, income.
31	**Stellman SD,** Djordjevic MV, Britton JA, et al. **Cancer Epidemiol Biomarkers Prev. 9(11):1241–9. 2000.**	case-control study **USA**	PCBs	Adipose tissue	After diagnosis	1994 – 1996	PCB 183 Concentration in adipose tissue (ng/g) OR 2.00; 95% CI (1.20 – 3.40).	Only PCB congener 183 is significantly associated with risk, with an adjusted odds ratio of 2.0 (95% confidence interval, 1.2–3.4) in women with adipose levels >5.67 ng/g;	Age, BMI, lipid, hospital, race.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
32	**Aronson KJ,** Miller AB, Woolcott CG, et al. **Cancer Epidemiol Biomarkers Prev.9(1):55–63. 2000.**	case-control study **Canada**	Breast adipose tissue concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk.	Adipose tissue	Before diagnosis	1995 – 1997	*Aroclor 1260 OR 1.15; 95% CI (0.58–2.25). *Total PCBs was calculated as the sum of PCBs 138 and 153 multiplied by 5.2 to approximate a level of the commercial PCB mixture, Alaclor 1260. PCB 105 = OR 3.17; 95% CI (1.51–6.68). PCB 118 = OR 2.31; 95% CI (1.11–4.76).	This study demonstrates increased breast cancer risk associated with the breast adipose tissue concentrations of some specific PCB congeners, but inconsistent or null results for the other compounds that were measured.	Age, BMI, race, lipid, family history, fat intake, alcohol.
33	**Muscat JE,** Britton JA, Djordjevic MV, et al. **Cancer Epidemiol Biomarkers Prev. 12(12):1474-8. 2003.**	hospital-based case-control study Long Island, New York . **USA**	PCBs, p,p -DDE, trans-nonachlor, oxychlordane, - hexachlorocyclohexane and hexachlorobenzene.	Adipose tissue	After diagnosis	1994 – 1996	RR of breast cancer recurrence associated with adipose concentrations of PCBs Total PCB Highest tertile OR 2.9; 95% CI (1.02–8.2) PCB 118 Highest tertile OR 4.0; 95% CI (1.32–4.9) PCB 118 Highest tertile OR 2.6; 95% CI (1.0–7.1) PCB 167 Highest tertile OR 3.1; 95% CI (1.0–9.3)	In summary, these results suggest that relatively high PCBs concentrations were related to increased rate of breast cancer recurrence, although this contrasts with our previous data showing no effect of OCs on breast cancer risk in these women.	Age, BMI, and stage of disease at diagnosis, spearman correlation coefficients were calculated to determine the associations between log OC concentrations and
34	**Wielsoe M,** Kern P, Bonefeld-Jørgensen EC. **Environmental Health 16:56, 2017;**	Case-control study of Inuit women from **Greenland**	Serum levels of environmental pollutants is a risk factor for breast cancer in Inuit: a case control study.	Serum	After diagnosis	2000-2003 and 2011-2014	Odds ratio of breast cancer risk associated with serum levels of PCB Grp2 (µg/Kg lipid) ΣPCB Grp2 (OR (95% CI)) 1.00 (reference) 2nd Tertile 2.28 (1.01; 5.18) 3rd Tertile 2.14 (0.94; 4.88)	Significant, positive associations between breast cancer risk and PCBs and PFAAs were observed. The associations indicate that environmental exposure to POPs can be a factor increasing the risk for breast cancer in Inuit women.	Information about age, BMI, smoking status, menopause status, number of full term pregnancies, and history of breastfeeding was obtained from questionnaires.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
35	**Holmes AK,** Koller KR et al. **International Journal of Circumpolar Health 73, 257-60. 2014;**	case-control study. Alaska **USA**	concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk.	Blood serum and Adipose tissue samples	After diagnosis	1999 - 2002	Persistent pesticides, PCBs, and most phthalate metabolites were not associated with case status in univariate logistic regression.	Our study is limited by small sample size and an inability to control for the confounding effects of body mass index.	Age, BMI, family history, lactation, menopausal status, hormone use, number of live birth.
36	**Liljegren G,** Hardell L, Lindström G, Dahl P, Magnuson A. **Eur J Cancer Prev. 1998.**	Case-control study. **Sweden**	PCBs, DDE, HCB.	Adipose tissue	After diagnosis	1993–1995	PCB, pg/g lipid PCB #77 > 4.5 pg/g lipid (OR = 5.8, 95% confidence interval (CI) = 0.8-42), PCB #126 > 145 pg/g lipid (OR = 2.2, 95% CI = 0.2-18), PCB #169 > 90 pg/g lipid (OR = 7.8, 95% CI = 0.6-96), and for HCB > 40 ng/g lipid (OR = 1.9, 95% CI = 0.4-7.2)	No significant difference for the sum of non co-planar PCBs or DDE was found between cases and controls.	Age, family history, menopausal status,
37	**Lucena RA**, Allam MF et al. Eur. J. Cancer Prev. 10, 117–119. **2001;**	Case-control study. **Spain**	Breast cancer risk factors: PCB congeners.	Adipose tissue	After diagnosis	1997	The most important risk factor identified by the model was PCB n-28 OR 9.6 (95% CI = 3.8 - 24.4)	The magnitude of association between PCB n-28 and breast cancer OR 9.6. together with the fact that its exposure proceeded the effect show the internal validity.	Age, family history, lactation, smoking habitat
38	**Hoyer AP,** Jørgensen T, Grandjean P, Hartvig HB. **Cancer Causes Control, 11(2):177–84. 2000;**	Copenhagen City Heart Study (CCHS). **Denmark**	Breast cancer risk factors: PCB congeners.	Serum	After diagnosis	1976 - 1983	A statistically significant increased risk and trend was found for subjects in the highest quartile of PCB-138 concentration (average of two measurements; OR, 2.1; 95% CI, 1.0–4.4; P for trend = 0.04). Elevated odds ratios were reported for the highest quartile of exposure to total PCBs and congeners PCB-118 and PCB-153 (OR, 1.6, 1.9 and 1.3, respectively	This study provides new evidence of the adverse effect of some organochlorines on breast cancer risk. Furthermore, repeated assessment of exposure during a relevant time period may provide a more precise risk estimate than a single measurement.	Age, BMI, height, number of full-term pregnancies, alcohol consumption, smoking, physical activity, menopausal status, household income, marital status, and education.

WASABY

ANNEX 1 – References PCBs monographs IARC, reviews, meta – analysis.

N°	Monographs IARC - PCBs	N°	Reviews - PCBs
1	**IARC Monographs PCB N° 107 - Lyon 2016**	1	**Mouly TA,** Toms LML. Environ Sci Pollut Res.Volume 23, Issue 22, pp 22385–22407. **2016.**
	Meta - Analysis - PCBs	2	**Brody JB**, **Cancer Supplement.** Volume 109 / Number 12. **2007.**
1	**Zhang J**, Huang Y, Wang X, et al. **PLOS ONE,** 10, **2015.**	3	**Wolff MS**, Toniolo PG. 103 (Suppl 7) :141-145 **EHP. 1995.**
2	**Leng L, Li J, et al. E**nvironment International 88. 133–141, **2016.**	4	**Rodgers KM,** Udesky JO, Rudel RA, Brody JG. Environmental Research 160,152–182. **2018.**

ANNEX 2 – References DDT, DDD, DDE and Organochlorines Compounds (Lindane, Hexachlorobenzene (HCB), Chlordane, ecc.) articles

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**López-Carrillo L**, Blair A, López-Cervantes M, Cebrián M, Rueda C, Reyes R et al. **Cancer Res, 57(17):3728–32. 1997.**	hospital-based case-control study **Mexico**	DDE (ng/g lipid) 505.46; p,p'-DDT(ng/g lipid) 84.53	Serum	After diagnosis	1994-1996	Not statistically significance	These results do not lend support to the hypothesis that DDT is causally related to breast cancer-	Age, BMI, socio economic characteristics, diet, and occupation.
2	**Romieu I,** Hernandez-Avila M, Lazcano-Ponce E, Weber JP, Dewailly E. Am J Epidemiol, 152(4):363–70. **2000.**	case-control study **Mexico**	DDE	Serum	After diagnosis	1990-1995	After adjustment for age, age at menarche, duration of lactation, Quetelet index, and serum DDT levels, serum DDE levels were positively related to the risk of breast cancer.	The data suggest that high levels of exposure to DDE may increase women's risk of breast cancer, particularly among postmenopausal women.	Age, menarche, lactation, BMI, menopause
3	**Boada LD,** Zumbado M, Henríquez-Hernández LA, Almeida-González M, Alvarez-León EE, Serra-Majem L et al. Environ Health, 11(1):28. **2012.**	case-control study. Gran Canaria Island, **Spain**	Complex organochlorine pesticides	Serum	After diagnosis	1999-2001	OR=1.097; 95% CI 0.420–28.412	Limitations: not matched by age; cases were significantly older than controls, and few women were exposed	Age, BMI, menopausal status, lactation, and tobacco
4	**Charlier C**, Albert A, Herman P, Hamoir E, et al. **Occup Environ Med, 60(5):348–51. 2003.**	case control study. Liege, **Belgium**	DDT and hexachlorobenzene (HCB)	Serum	After diagnosis	/	Not statistically significance	Mean levels of total DDT and HCB were significantly higher for breast cancer patients than for controls but Not statistically significance	Age, BMI, menopausal status, blood levels of HCB and total DDT
5	**Charlier C**, Foidart JM, Pitance F, et al. **Clin Chem Lab Med, 42(2):222–7. 2004.**	case control study. Liege, **Belgium**	DDE and hexachloro benzene HCB	Serum	After diagnosis	2001 - 2002	p,p DDE > 0.5 ppb OR 2.21 95% CI (1.41–3.48) HCB> 0.5 ppb OR 4.99, 95% CI (2.95-8.43)	Strengths: sample size; blood draw before surgery Limitations: no BMI information; hospital based controls	Age, BMI, parity, lactation, menopause, HRT, family history

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
6	**Zheng T**, Holford TR, Mayne ST et al. **Cancer Epidemiol Biomarkers Prev, 9(2):167–74. (2000).**	Case–control study. Connecticut, **USA**	DDE and PCBs	Serum Samples	After diagnosis	1995–1997	DDE (highest versus lowest quintile adjusted for age, body mass index (BMI) and breastfeeding for DDE: OR = 1.02, 95% CI = (0.61, 1.72), p-trend = 0.74;	No association was identified between DDE and risk of breast cancer	Age, BMI, reproductive risk factors, HRT, dietary fat intake, family history of breast cancer.
7	**Itoh H**, Iwasaki M, Hanaoka T et al. Cancer Causes Control, 20(5):567–80. **2009.**	Case–control hospital study in Nagano Prefecture. **Japan**	DDT, DDE	Serum Samples	After diagnosis	2001 - 2005	IV Quartile median (ng/g lipid) p,p′-DDE:1.02 (0.46–2.26) p,p′-DDT:0.58 (0.27–1.25) o,p′-DDT:0.67 (0.30–1.50)	No association was identified between DDT, DDE and risk of breast cancer	Age, BMI, total lipid concentration in serum, reproductive risk factors, medical history, area, lipid, menopause, smoking, fish and veg consumption, family history of breast cancer
8	**Wolff MS**, Toniolo PG, Lee RW, et al. **J Natl Cancer Inst 85:648–52.1993.**	A prospective cohort study. New York. **USA**	DDE, PCBs	Serum Samples	After diagnosis	1985 - 1991	Increase in relative risk of breast cancer for an elevation of serum DDE concentrations from 2.0 ng/mL (10th percentile) to 19.1 ng/mL (90th percentile).	In this population of New York City women, breast cancer was strongly associated with DDE in serum but not with PCBs.	Age, BMI, after adjustment for first-degree family history of breast cancer, lifetime lactation, and age at first full-term pregnancy.
9	**Gatto NM**, Longnecker MP, Press MF et al. **Causes Control.18:29–39. 2007.**	Case–control study in five **USA**	DDE, DDT, PCBs.	Blood Samples	After diagnosis	1995-1998	DDE: OR = 1.02, 95% CI = (0.61, 1.72), p-trend = 0.74; PCBs: OR = 1.01, 95% CI = (0.63,1.63),p-trend = 0.56).	No association was identified between DDT, DDE, PCBs and risk of breast cancer	Age, BMI, lactation, lipid.
10	**Millikan R**, DeVoto E, Duell EJ, et al. **Cancer Epidemiol Biomarkers Prev. 9(11):1233–40. 2000.**	case-control study **USA**	DDE, PCBs	Plasma	After diagnosis	1993 – 1996	DDE, µg/g lipid African-Americans ≥ 1.8 OR 1.41; 95% CI (0.87–2.29)	We observed no overall association between plasma levels of DDE and total PCBs and breast cancer among African-American and white women in North Carolina.	Age, BMI, age-squared, race (all participants), lipid, menopausal status, parity/lactation, HRT, income.
11	**Gammon MD,** Wolff MS, et al. Vol. 11, 686–697, **Cancer Epidemiology, Biomarkers & Prevention; 2002;**	Case-control study. Long Island, **USA**	DDE, DDT, Chlordane, Dieldrin, PCBs	Blood Samples	After diagnosis	1996-1997	No substantial elevation in breast cancer risk was observed in relation to the highest quintile of lipid-adjusted serum levels of organochlorine compounds	No association was identified between DDE, DDT, Chlordane, Dieldrin, PCBs and risk of breast cancer	Age, BMI, race, reproductive history, benign breast disease Interview response rates: cases, 83.2%; controls, 68.0%.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
12	**Demers A**, Ayotte P, Brisson J et al.**Cancer Epidemiol Biomarkers Prev, 9(2):161–6. 2000.**	Case-control study. **Canada**	DDE, PCBs.	Serum Samples	After diagnosis	1994–1997	DDE, ng/g lipid, population controls 282.5 < 427.8 Population OR (95% CI) 1.06 (0.62–1.79) Hospital OR (95% CI) 0.66 (0.37–1.19)	No association was identified between DDE, PCBs and risk of breast cancer	Age, region of residence, BMI, breast feeding duration, age at first child, number of fertile years, family history of breast cancer, and history of benign breast disease.
13	**Laden F**, Hankinson SE, Wolff MS et al. **Int J Cancer.** 91(4):568–74. **2001.**	nested case-control study. **USA**	DDE and PCBs	Plasma	After and before diagnosis	1976–1994	DDE, µg/g lipid 1.466 – 6.054 OR (95% CI) 0.82 (0.49–1.37)	No association was identified between DDE, PCBs and risk of breast cancer	Age, BMI, family history of breast cancer, history of BBD, age at menarche, BMI, lipid, number of children, age at birth of first child, lactation.
14	**Krieger N,** Wolff MS, Hiatt RA, et al. **J Natl Cancer Inst, 86(8):589–99. 1994.**	5 small case–control study. San Francisco Bay **USA**	DDE, DDT and PCBs.	Serum Samples	after (or at most 6 months before) diagnosis.	1964 - 1990	no significant association was seen between risk of cancer of the breast and serum DDE	The data do not support the hypothesis that exposure to DDE and PCBs increases risk of breast cancer.	BMI, age at menarche, ever vs never pregnant, menopausal status at time of case patient´s diagnosis of breast cancer plus variables matched by design.
15	**Hunter DJ,** Hankinson SE, Colditz GA,et al. **N Engl J Med, 337(18):1253–8. 1997.**	case-control study **USA**	DDE	Plasma	After diagnosis	1976 - 1992	DDE, ppb > 9.46 (OR, 0.72; 95% CI, 0.37–1.4).	Our data do not support the hypothesis that exposure to DDE increases the risk of breast cancer.	History of breast cancer in a mother or sister of benign breast disease, age at menarche, number of children and age at birth of first child, duration of lactation, BMI, plus variables matched by design
16	**Dorgan JF**, Brock JW, Rothman N et al. **Cancer Causes Control. 10(1):1–11. 1999.**	nested case-control study **USA**	DDT, PCBs	Blood samples	before diagnosis	1977 - 1987	No association found for DDE and DDT in ng/g lipid in serum of exposed cases	Results of this study do not support a role for organochlorine pesticides and PCBs in breast cancer etiology.	Age, BMI, menopausal status, family history, estrogen use, parity, age at menarche

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
17	**Helzlsouer KJ,** Alberg AJ, Huang HY, et al. **Cancer Epidemiol Biomarkers Prev, 8(6):525–32. 1999.**	Nested case-control study. **USA**	DDE, DDT, PCBs	Serum	Strengths: serum collected up to 20 yrs before diagnosis;	1974 - 1994	No association found for DDE in ng/g lipid in serum of e exposure category or level (1974 and 1989)	Even after 20 years of follow-up, exposure to relatively high concentrations of DDE or PCBs showed no evidence of contributing to an increased risk of breast cancer.	Age, history of breast cancer, BMI at age 20 yrs or current age at menarche, age at first birth, duration of lactation plus matching variables
18	**Cohn BA, EHP, 2007.**	Nested case–control study. Oakland **California**	DDT, DDE, and metabolites	Blood Samples (young women) mean age of 26 years	Before diagnosis	1959–1967	Age < 14 yrs in 1945 p,p′-DDT, µg/L 8.09–13.90 (OR, 2.5; 95% CI,1.0–6.3)	Exposure to p,p´-DDT early in life may increase breast cancer risk.	Age (matching), BMI, blood lipids (total cholesterol, total triglycerides), parity, year of blood draw, breast-feeding after current pregnancy 10 congeners measured.
19	**Cohn BA,** La Merrill M, Krigbaum NY et al. **J Clin Endocrinol Metab, 100(8):2865–72. 2015.**	Nested case–control study. Oakland. **California**	DDT	Serum post partum	Before diagnosis	1959–1967	Maternal o,p -DDT predicted daughters' breast cancer (odds ratio fourth quartile vs first (OR, 3.7; 95% CI, 1.5–9.0).	This prospective human study links measured DDT exposure in utero to risk of breast cancer.	Age, BMI, maternal cholesterol and triglycerides, maternal overweight in early pregnancy and maternal history of breast cancer
20	**Hoyer AP,** Jørgensen T, Grandjean P et al. **Cancer Causes Control, 11(2):177–84. 2000.**	nested case-control study Copenhagen. **Denmark**	DDT	Serum	After diagnosis	1976 - 1992	p,p' DDT ng/g OR= 3.6; 95% CI, 1.1–12.2)	A high serum concentration of p,p'-DDT over the course of the two examinations was associated with a more than three-fold significantly increased risk of breast cancer.	Age, BMI, breast cancer characteristics included tumor size, degree of spread, and stage of disease.
21	**Ward EM,** Schulte P, Grajewski B et al. **Cancer Epidemiol Biomarkers Prev. 9(12):1357-67. 2000.**	nested case-control study **Norway**	DDE and PCBs	Serum	before diagnosis	1973 - 1991	no significant association was seen between risk of cancer of the breast and serum DDE	This is a case-control study of serum organochlorine levels in relation to breast cancer risk.	Age, BMI, occupational category, age at first birth, number of births prior to donation of blood sample, region of residence, and region of birth, lipid concentration in serum.
22	**Iwasaki M**, Inoue M, Sasazuki S, et al. **Science Total Environment 402(2–3):176–183. 2008.**	Nested case–control study. **Japan**	DDT, DDE, HCB, β-HCH.	Plasma	Before diagnosis	1990 - 1994	Adjusted ORs for p,p'-DDT, HCB,and β-HCH were less than 1. For p,p'-DDE, adjusted OR for the highest versus lowest quartile was 1.48 (95% confidence interval 0.70-3.13; p for trend=0.25).	We found no statistically significant positive association between plasma organochlorine level and breast cancer risk.	Age at menarche, menopausal status at baseline, number of births, age at first birth, height (continuous), BMI, alcohol consumption

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
23	**Raaschou-Nielsen O,** Pavuk M, LeBlanc A, et al. **Cancer Epidemiol Biomarkers Prev. 14(1):67–74. 2005.**	Nested case-control study **Denmark**	DDE	Adipose breast tissue	Tissue was sampled up to 6.5 years before the breast cancer diagnosis	1993 – 1997	No association found for DDE and DDT in ng/g lipid in serum of exposed cases	We found no indication of higher breast cancer risk in association with higher adipose tissue concentrations of any of the chlorinated pesticides.	Age, education, BMI, lipid, alcohol, number of children, age at birth of first child, lactation, HRT, history of BBD.
24	**Wolff MS**, Berkowitz GS, Brower S et al. **Environ Res, 84(2):151–61. 2000b**	A hospital-based case-control study. NY **USA**	DDE, DDT	blood serum or adipose tissue	After diagnosis	1985 - 1994	DDE: 664–1172 ng/g OR= 0.81;95% CI,0.35–1.87	We found no statistically significant positive association between plasma organochlorine level and breast cancer risk.	Age at menarche, number of full-term pregnancies, age at first full-term pregnancy, family history of breast cancer, lifetime history of lactation, height), BMI, BMI-menopausal status interaction
25	**Olaya-Contreras P,** Rodríguez-Villamil J, Posso-Valencia HJ, et al. **Cad Saude Publica, 14:Suppl 3: 125–32. 1998.**	Bogota **Colombia**	DDT, DDD and DDE	Serum	After diagnosis	1995–1996	DDE, ng/mL (higher category of DDE exposure) 1.97–19.20 OR=1.95; 95% CI,1.10–3.52	We confirm that serum DDE levels bear a positive association to risk of breast cancer	Age, family history, BMI, parity, menopause, breast cancer history, lactation
26	**Moysich KB**, Ambrosone CB, Vena JE, et al. **Cancer Epidemiol Biomarkers Prev. 7(3):181–8. 1998.**	case-control study **USA**	DDE, HCB,PCBs	Serum	After diagnosis	1986-1991	DDE, ng/g lipid 3rd tertile OR=1.34; 95% CI, 0.71–2.55	Elevated serum levels od DDE were not associated with breast cancer risk	Age, education, family history of breast cancer, parity, quetelet index, lactation, age at first birth, years since last pregnancy, fruit and vegetable intake, lipid.
27	**Mendonca GA**, Eluf-Neto J, Andrada-Serpa MJ, et al. **Int J Cancer, 3(5):596–600. 1999.**	hospital-based case-control study. **Brasil**	DDE	Serum	After diagnosis	1995–1996	Serum DDE, ng/mL ≥ 7.6 OR=0.83; 95% CI, 0.4–1.60	Exposure to organochlorinated pesticides measured by history or serum analysis was thus not a risk factor for breast cancer.	Age, lactation, education, parity, smoking, family history, breast size
28	**Dello Iacovo R**, Celentano E, Strollo AM, et al. **Adv Exp Med Biol, 472:57–66. 1999.**	case-control study. Naples **Italy**	DDE	Serum	After diagnosis	1993-1998	Serum DDE, ng/mL > 10.2 OR=1.24; 95% CI, 0.7–2.2	Strengths: sizeable sample; > 30% having DDE > 10 ng/mL	Age, BMI, lactation, parity, serum lipids, education, smoking, menopause
29	**Wolff MS**, Zeleniuch-Jacquotte A, Dubin N, et al. **Cancer Epidemiol Biomarkers Prev, 9(3):271–7. 2000a**	hospital case-control study **USA**	DDE and PCBs	Serum	Before diagnosis	1994 - 1995	DDE, µg/g lipid 0.034–1.3 OR=1.34; 95% CI, 0.82–2.20	In summary, we found no association between organochlorine compounds and breast cancer	Age, age-square, menopause, race, strata, lactation, HRT, parity

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
30	**Schecter A**, Toniolo P, Dai LC, Thuy LT, et al. **Arch Environ Contam Toxicol, 33(4):453–6. 1997.**	Small hospital-based case-control study. Hanoi. **Vietnam**	DDT and DDE	Serum	After diagnosis	1994	DDE, ng/mL 3rd vs 1st tertile OR =1.21; 95% CI, 0.15–9.65	levels and breast cancer risk in our study.	Age, BMI, age at menarche, parity, lactation, weight
31	**Soliman AS**, Wang X, Di Giovanni J, Eissa S, et al. **Environ Res, 92(2):110–7. 2003.**	case-control study **Egypt**	DDE and beta - HCH	Serum	After diagnosis	N. R.	DDE (ppb) > 4.7 OR =1.41; 95% CI, 0.63–3.19	Organochlorine serum level was not a risk factor of breast cancer in this population.	Age, lactation, menopausal status, residence.
32	**Pavuk M**, Cerhan JR, Lynch CF, et al. **J Expo Anal. Environ Epidemiol, 13(4):267–75. 2003.**	case-control study. **Slovakia**	**DDE, DDT,** and hexachlorobenzene (HCB) andaA total of 15 individual **PCB congeners.**	Serum	After diagnosis	1997–1999	DDE, ng/g lipid 4389–19912 OR=3.04; 95% CI, 0.65–14.3 DDT, ng/g lipid 137–562 OR=1.19; 95% CI, 0.27–5.23	Organochlorine compounds was positively associated with risk, but not statistically significant	Age, menarche, education, alcohol intake, smoking.
33	**Rubin CH**, Lanier A, Kieszak S, et al. **International journal of circumpolar health** 65(1):18–27. **2006.**	Case-control study; **Alaska**	DDT, DDE and 13 other chlorinated pesticides, PCBs (28 congeners).	Serum Samples	Before diagnosis	1983-1987	DDE (ppb) > 9.62 OR=1.43; 95% CI, 0.46–4.47	we found no statistically significant association between organoclorine compounds and breast cancer	Parity, family history of breast cancer, race, triglycerides, cholesterol.
34	**van't Veer P**, Lobbezoo IE, Martín-Moreno JM, et al. **BMJ, 315(7100):81–5. 1997.**	multicentre case-control study. **Germany, the Netherlands, Ireland, Switzerland, and Spain**	DDE, DDT	Adipose tissue	After diagnosis	1991–1992	DDE, µg/g 0.87- 1.89 OR= 1.14; 95% CI, 0.62–2.12	This study does not support the hypothesis that DDE increases risk of breast cancer in postmenopausal women in Europe.	Age, centre, BMI, age at first birth, alcohol consumption
35	**Liljegren G**, Hardell L, Lindström G, et al.. **Eur J Cancer Prev. 1998.**	Case-control study. **Sweden**	DDE, HCB and PCBs	Adipose tissue	After diagnosis	1993–1995	DDE, ng/g lipid > 700 OR=0.4; 95% CI, 0.1–1.2	No significant difference for the sum of non co-planar DDE or PCBs was found between cases and controls.	Age,BMI, menopausal status, parity.
36	**Zheng T**, Holford TR, Mayne ST, et al. Cancer, 85(10):2212–8.1999.	case-control study Connecticut, **USA.**	DDE and DDT	Adipose tissue	After diagnosis	1994–1997	DDE, ng/g lipid 412.6–779.2 OR=1.3; 95% CI, 0.7–2.2	These results do not support an association between adipose tissue levels of DDE and DDT and breast cancer risk.	Age, BMI, lifetime months of lactation, age at menarche, age at FFTP, menopausal status.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
37	Aronson KJ, Miller AB, Woolcott CG, et al. **Cancer Epidemiol Biomarkers Prev.9(1):55–63. 2000.**	case-control study **Canada**	DDE and PCBs	Adipose tissue	Before diagnosis	1995 – 1997	DDE, µg/kg lipid > 1390 OR=1.62; 95% CI, 0.84–3.11	These results do not support an association between adipose tissue levels of DDE and PCBs and breast cancer risk.	Age, study site, menopausal status, present use of HRT, ethnicity, BMI fat and alcohol intake
38	Ibarluzea JM, Fernández MF, Santa-Marina L, et al. **Cancer Causes Control, 15(6):591–600. 2004.**	Hospital case-control study. **Spain**	DDE, aldrin, endosulfan and lindane	Adipose tissue	After diagnosis	1996–1998	DDE, ng/g lipid ≥ 675.98 OR, 1.22; 95% CI, 0.68–2.21	DDE was positively associated with risk, but not statistically significant	Age, reference hospital, number of children, age at FFTP, family history of breast cancer, and alcohol and tobacco consumption
39	**Hoyer AP**, Jorgensen T, Rank F, et al. **BMC Cancer. 1:8. 2001.**	Nested case-control study. **Denmark**	Dieldrin, HCB, PCBs	Blood Samples	Before diagnosis	1976 - 1978	Organochlorines in quartiles (ng/mL) Dieldrin >57.11 Estrogen receptor negative OR=7.6; 95% CI,1.3-46.1	we observed increased breast cancer risk associated with exposure to dieldrin derived from women who developed an estrogen receptor negative (ERN) tumor	Breast cancer characteristics included tumor size, degree of spread, and stage of disease
40	Hoyer AP, Grandjean P, Jørgensen T et al. **The Lancet 352(9143): 1816 – 20.1998.**	nested case-control study **Denmark**	DDT, Dieldrin, PCBs	Serum	After diagnosis	1976 – 1995	DDT OR=0·84; 95% CI,0·49–1·45	Our results support the hypothesis that organochlorine compounds such as dieldrin, which have oestrogenic properties, may increase the risk of breast cancer.	age, weight, height, number of children, alcohol, smoking, lipid, physical activity, menopausal status, household income, marital status, education.

ANNEX 2 – References DDT, DDD, DDE and Organochlorines Compounds (Lindane, Hexachlorobenzene (HCB), Chlordane, ecc.) monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - DDT	N°	REVIEWS - DDT, DDD, DDE and Organochlorine Compounds
1	**IARC Monographs DDT, LINDANE, AND 2,4-D Vol. 113; 2018**	1	**López-Cervantes M et al. Environ Health Perspect. 2004 Feb;112(2):207-14.**
	Meta - Analysis - DDT, DDD, DDE	2	**Wolff MS**, Toniolo PG. 103 (Suppl 7) :141-145 **EHP. 1995.**
1	**Park JH,** Cha ES, Ko Y, et al. Osong Public Health Res Perspect. Apr;5(2):77-84. **2014.**	3	**Snedeker SM**, **EHP. 2001.**
		4	**Gray JM** , Rasanayagam S, Engel C and Rizzo J. **Environmental Health, 16:94; 2017.**
		5	**Calle EE,** Frumkin H, Henley SJ, et al. **CA Cancer J Clin. Sep-Oct;52(5):301-9. 2002.**
		6	**Macon MB**, Fenton SE. J Mammary Gland Biol Neoplasia. Mar;18(1):43-61. **2013.**
		7	**Gray JM** , Rasanayagam S, Engel C and Rizzo J. Environmental Health, 16:94; **2017.**

ANNEX 3 – References Dioxins (TCDD) articles.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**Warner M**, Mocarelli P, Samuels S, **Environ Health Perspect. Dec; 119(12): 1700-5. 2011.**	Cohort study. Seveso (MI). **Italy**	TCDD	Blood draw	/	1996 - 2008	For breast cancer, the HR was increased, but not significantly (adjusted HR = 1.44; 95% CI: 0.89, 2.33).	Individual serum TCDD is significantly positively related with all cancer incidence in the SWHS cohort, more than 30 years later. This all-female study adds to the epidemiologic evidence that TCDD is a multisite carcinogen.	Individual TCDD concentration was measured in archived serum by high-resolution mass spectrometry. A total of 833 women participated in the 2008 follow-up study. We examined the relation of serum TCDD with cancer incidence
2	**Pesatori AC**, Consonni D, Rubagotti M, Grillo P, Bertazzi PA. **Environ Health. 8:39. 2009.**	Cohort study. Seveso (MI). **Italy**	2,3,7,8-tetrachloro dibenzo-p-dioxin (TCDD);	Blood samples	/	1977-1996	An increased risk of breast cancer was detected in zone A females after 15 years since the accident (five cases, RR, 2.57; 95% CI, 1.07-6.20).	The elevated risk of breast cancer in zone A females after 15 years since the accident deserves further and thorough investigation.	Gender-, age-, and period-adjusted rate ratios (RR) and 95% confidence intervals (95% CI) were calculated .
3	**Bencko V**, Rames J, Ondrusova M, at al. **Neoplasma. 2009;56(4):353-7. 2009.**	Two ecological studies in **Slovak Republic and in Czech Republic.**	TCDD and PCBs	/	/	1987-1996	Neither PCBs nor TCDDs/Fs appear to contribute to the observed significantly lower incidence of breast and prostate cancer in the Michalovce District and lower bladder cancer incidence in Uherske Hradiste District.	The results of our analysis points to substantial potential problems of risk assessment for cancer incidence in populations exposed to xenobiotic, or more generally, as it relates to a wide spectrum of confounding of cancer risk factors.	The age-adjusted world standard ratio (WSR) incidence of malignant breast tumors in females were compared for whole the Slovak Republic and the Michalovce District. For
4	**Warner M**, Eskenazi B, Mocarelli P, et al. Environ Health Perspect. Jul;110(7):625-8. **2002.**	Cohort study. Seveso (MI). **Italy**	TCDD	serum samples	before and after diagnosis	1996-1998	We observed a statistically significant, dose response–increased risk for breast cancer incidence with individual serum TCDD level among women in the Seveso Women's Health Study.	Continued follow-up of the cohort will help shed light on the possible role of TCDD in the pathogenesis of breast cancer	We examined the effect of potential confounders like gravidity, parity, age at first pregnancy, age at last pregnancy, lactation, family history of breast cancer, age at menarche, current body mass index.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
5	**Revich B**, Aksel E, Ushakova T, et al. **Chemosphere. May - Jun;43(4-7):951-66. 2001.**	Ecological study. Chapaevsk, **Russia.**	Dioxins;	blood samples and human milk	/	1997-1998	The SMR is higher for breast cancer. Chapaevsk women have a higher risk overall due to breast cancer 2.1 (C.I. 1.6-2.7)	We confirm that blood dioxins levels bear a positive association to risk of breast cancer	Sampling of human milk was carried out. The average age of the mothers was 22.0 years. All surveyed mothers classified their diet as mixed and diary products every day. All respondents never smoked as they self-reported in the interview.
6	**Bertazzi PA,** Zocchetti C, Guercilena S, Consonni D, Tironi A, Landi MT, Pesatori AC. **Epidemiology**. Nov;8(6):646-52. **1997.**	A 15-year mortality study. **Italy**	2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD).	/	/	1976-1991	We found no increase for all-cancer mortality or major specific sites (for example, respiratory among males, breast among females).		age, BMI, smoking, lipid, physical activity, menopausal status.
7	**Bertazzi PA**, Pesatori AC, Consonni D, Tironi A, Landi MT, Zocchetti C. Vol. 4, No. 5, pp. 398-406. **Epidemiology. 1993.**	Cohort study. Seveso (MI). **Italy**	2,3,7,8-Tetrachlorodibenzo-para-Dioxin (TCDD).	serum	/	1977-1986.	Breast cancer among females was below expectations in the most contaminated zones.		age, BMI, smoking, lipid, physical activity, menopausal status.
8	**Viel JF**, Clément MC, Hägi M, Grandjean S, Challier B, Danzon A. **Int J Health Geogr. 28;7:4. 2008.**	Case-control study. **France**	Dioxins	/	/	1996 - 2002	no increased or decreased risk was found for any dioxin exposure category. Conversely, women over 60 years old living in the highest exposed zone were 0.31 time less likely (95% confidence interval, 0.08–0.89) to develop invasive breast cancer.	Before speculating that this decreased risk reflects a dioxin anti-estrogenic activity with greater effect on late-onset acquired breast cancer, some residual confounding must be envisaged.	age, BMI, smoking, lipid, physical activity, menopausal status.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
9	**Reynolds P**, Hurley SE, Petreas M, et al. **Cancer Causes Control; 16(5):525-35. 2005.**	A hospital-based case-control study. In the San Francisco Bay Area. **USA**	PCDD/PCDFs.	Breast adipose tissue	at (or near) the time of diagnosis.	during the mid-1990s	One notable exception was octachlorodibenzo-p-dioxin (OCDD), for which the odds ratio for the second and third tertiles appeared modestly elevated (OR = 1.22, 95% CI: 0.47:3.16 and OR = 1.62, 95% CI: 0.64:4.12, respectively), though the test for trend was not significant (p = 0.36).	Breast cancer risk was not associated with adipose levels of PCDD/PCDFs. More study is suggested among women of color who may have higher body burden levels of these compounds.	Invasive breast cancer cases in women ≥20 year of age were identified from the California Cancer Registry, for 1988-1997. Neighbourhood SES and urbanization were derived from U.S. Census data.
10	**Verkasalo PK**, Kokki E, Pukkala E, et al. **Environ Health Perspect. 112(9):1026–1031. 2004.**	The ecological study population. **Finland**	PCDD/Fs	/	/	1981 - 2000	RR 1.15 (1.03 – 1.28 95% C.I.) in those living 1.0 – 4.9 Km from the river.	We also observed increases for cancers of the uterine cervix and corpus, breast, and lung, and BCCs among those living 1.0 – 4.9 km from the river.	All variables were classified according in 500 m × 500 m grid squares. For each grid square in Finland, observed cancers were counted by sex, age, and SES.
11	**Hardell L**, Lindström G, Liljegren G, et al. **Eur J Cancer Prev. Oct;5(5):351-7. 1996.**	Case control study. **Sweden**	PCDD/Fs	Breast adipose tissue	After diagnosis	1993 - 1995	OR was obtained for OCDD: 401-1000 pg/g lipid yielded OR 3.8, 95% confidence interval (CI) 0.4-39	Breast tissue concentration of OCDD was increased in cancer patients but not statistically significant, whereas the concentrations of other PCDDs and PCDFs were equal in cases and controls.	age, BMI, smoking, breast adipose tissue
12	**Dai D,** Oyana TJ. **Environ Health. 2008.**	Ecological study. Michigan, **USA.**	Dioxins	/	/	1985 - 2002	Preliminary statistical analysis suggests that there is a strong association between elevated levels of breast cancer incidence and aging, particularly among females residing in the city of Midland or near areas contaminated with high dioxins levels.	Increased breast cancer incidences are spatially associated with soil dioxin contamination.	GIS analysis supported; cancer registry in the MDCH; Each case includes information on patient's gender, ZIP code of a patient's residence, year of diagnosis, primary site, stage at diagnosis, and age group; No socio-economic status

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
13	**Guajardo OA and Oyana TJ. Journal of Environmental and Public Health Volume.** Article ID 316249, 16 pages. **2009.**	Ecological study. Michigan, **USA.**	Dioxins	/	/	1985 - 2002	ZIP codes 48640 OR = 1.76, (95% CI 1.316 – 2.355) ZIP codes 48603 OR = 1.65, (95% CI 1.238 – 2.202) ZIP codes 48734 OR = 1.88, (95% CI 1.349–2.630)	the spatial distribution of breast and lung cancer incidence rates is non homogeneous; a significant positive association between possible exposure to environmental pollution and risk of breast and lung cancer was found;	the socioeconomic factors, such as race and residency at the same location, are more likely to explain spatial variability of cancer incidences.

ANNEX 3 – References Dioxins (TCDD) monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - Dioxin (TCDD)	N°	REVIEWS – Dioxin (TCDD)
1	**IARC Monographs Dioxin (TCDD) Vol. 69; 1997**	1	**Laden F,** Hunter DJ. **Annu Rev Public Health. 19:101-23.1998.**
2	**IARC Monographs Dioxin (TCDD) Vol. 100F. 2012.**	2	**Birnbaum LS**, Fenton SE. **Environ Health Perspect. 111(4): 389–394. EHP. 2003.**
		3	**Boffetta P,** Mundt KA, Adami HO, et al. **Crit Rev Toxicol. 41(7):622–36. 2011.**
		4	**Jenkins S,** Betancourt AM, Wang J, Lamartiniere CA. **J Steroid Biochem Mol Biol. Apr;129 (3-5):191-200. 2012.**
		5	**Macon MB, J Mammary Gland Biol Neoplasia. 2013.**

ANNEX 4 – References Polycyclic Aromatic Hydrocarbons (PAH) articles.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**Shantakumar SJ**, Gammon MD, Eng SM, et al.. **J Expo Anal Environ Epidemiol. 15(6):482-90.2005.**	Population - based study. Long Island, New York. **USA**	PAH	blood samples	After diagnosis	1996-1997	Women who donated blood in summer and fall had increased odds of detectable PAH-DNA adducts OR = 2.65, 95% confidence interval (CI)=1.69, 4.17;	These data suggest that PAH-DNA adducts detected in a population-based sample of adult women with ambient exposure levels reflect some key residential PAH exposure sources assessed in this study, such as cigarette smoking.	Ambient PAH exposure at the current residence was estimated using geographic modeling Environmental home samples of dust and soil were collected on a random subset of long-term residents (15þ years). A questionnaire, including a dietary history; environmental home samples; and geographic modeling.
2	**Beyea J,** Hatch M, Stellman SD. **Environmental Health Perspectives. Vol. 114, N° 7 2006**	population-based study Long Island, New York. **USA**	PAHs	blood samples	/	1996 - 1997	This study indicates that in developing inhalation exposure estimates it is necessary to account for emissions at intersections to fully determine the spatial distribution of PAH exposure.	This study indicates that in developing inhalation exposure estimates it is necessary to account for emissions at intersections to fully determine the spatial distribution of PAH exposure.	We have constructed a geographic model for airborne polycyclic aromatic hydrocarbons (PAHs) from traffic that is being used in a population-based, case–control epidemiologic study involving about 3,000 women on Long Island, New York, known as the Long Island Breast Cancer Study Project

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
3	**Gammon MD,** Santella RM, Neugut AI, et al. **Cancer Epidemiol Biomarkers Prev 11(8): 677–685. 2002.**	Case-control study. in Long Island, New York. **USA**	PAH.	blood samples	After diagnosis	1996 - 1997	The age-adjusted odds ratio (OR) for breast cancer in relation to the highest quintile of adduct levels compared with the lowest was 1.51 [95% confidence interval (CI), 1.04-2.20],	These data indicate that PAH-DNA adduct formation may influence breast cancer development, although the association does not appear to be dose dependent and may have a threshold effect.	women include lower parity, late age at first birth, little or no breast feeding, and family history of breast cancer. Factors associated with an increased probability include white or other race, alcohol use, ever breastfed, ever use of hormone replacement therapy, ever use of oral contraceptives, and ever had a mammogram.
4	**Gammon MD**, Sagiv SK, Eng SM, et al. **Arch Environ Health, 59(12):640-9. 2004.**	Case-control study. in Long Island, New York. **USA**	PAH	blood samples	After diagnosis	1996 - 1997	The odds ratio for breast cancer was elevated in relation to detectable PAH-DNA adducts (1.29 as compared with non detectable adduct levels; 95% confidence interval = 1.05, 1.58).	These data indicate that PAH-DNA adduct formation may influence breast cancer development, although the association does not appear to be dose dependent and may have a threshold effect.	women include lower parity, late age at first birth, little or no breast feeding, and family history of breast cancer. Factors associated with an increased probability include white or other race, alcohol use, ever breastfed, ever use of hormone replacement therapy, ever use of oral contraceptives, and ever had a mammogram.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
5	**Terry MB**, Gammon MD, Zhang FF, et al. **Cancer Epidemiol Biomarkers Prev. 13:2053–2058. 2004.**	Case-control study. In Long Island, NY. **USA**	PAH	blood samples	After diagnosis	1996-1997	Overall, we found a modest, statistically significant association between those subjects with at least one variant Gln allele at exon 23 of the XPD gene and breast cancer risk (OR, 1.21; 95% CI, 1.01-1.44).	Overall, this study suggests that those individuals with this polymorphism in the XPD gene may face an increased risk of breast cancer from PAH-DNA adducts and cigarette smoking.	Respondents were asked about their pregnancy, occupational, and residential history; their use of pesticides in their home or on a farm; electrical appliance use; lifetime history of consumption of smoked or grilled foods; medical history; family history of cancer; body size changes by decade; recreational physical activities; cigarette smoking; alcohol use; menstrual history; use of exogenous hormones; and demographic characteristics
6	**Sagiv SK**, Gaudet MM, Eng SM, et al.. **Environ Res. 109(3): 287–291. 2009.**	A survival analysis In Long Island. **USA**	PAH	blood samples	After diagnosis	1996 - 1997	there are not evidence that all-cause mortality (HR) = 0.88; 95% confidence interval (CI): 0.57–1.37), or breast cancer mortality (HR = 1.20; 95% CI: 0.63–2.28) was strongly associated with detectable PAH-DNA	Results from this large population-based study do not provide strong support for an association between detectable PAH-DNA adducts and survival among women with BC, except perhaps among those receiving radiation treatment.	case-control questionnaire, including family history of breast cancer; body size; physical activity; menstrual and reproductive histories; exogenous hormone use; active and passive cigarette smoking; and alcohol consumption. Medical records were abstracted for tumor characteristics including estrogen/ progesterone receptor

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
7	**Burdick AD**, Davis II JW, Liu KJ, et al. **Cancer Research 63, 7825–7833. 2003.**	Study of human mammary epithelial cells and the human mammary epithelial cell line MCF-10A.	PAHs	/	/	2003	/	Taken together, these data indicate that BPQs, through the generation of hydrogen peroxide, activate the EGFR in MCF-10A cells, leading to increased cell number under EGF-deficient conditions.	Benzo(a)pyrene Quinones Increase Cell Proliferation, Generate Reactive Oxygen Species, and Transactivate the Epidermal Growth Factor Receptor in Breast Epithelial Cells
8	**Plıskova M**, Vondrácek J, Vojtesek B, et al. **Toxicological Sciences 83, 246–256. 2005.**	Study of deregulation of Cell Proliferation by Polycyclic Aromatic Hydrocarbons in Human Breast Carcinoma MCF-7 Cells	PAHs	/	/	2005	/	In the present study, we found that two PAHs, benz[a]anthracene (BaA) and BaP, can stimulate proliferation of human breast carcinoma MCF-7 cells at concentrations 100 nM and higher.	Deregulation of Cell Proliferation by Polycyclic Aromatic Hydrocarbons in Human Breast Carcinoma MCF-7 Cells Reflects Both Genotoxic and Nongenotoxic Events
9	**Bonner MR,** Han D, Nie J, et al. **Cancer Epidemiol Biomarkers Prev 2005;14(1). 2005.**	Case -control study. In Western New York. **USA**	PAHs and TSP	/	/	1996-2001	For risk associated with estimated residential TSP concentrations on a continuous scale, in postmenopausal women, we observed a 20% increase in the odds ratio for every 30 μg/m³ increase in TSP concentration (adjusted OR, 1.20; 95% CI, 1.04-1.38).	Our study suggests that exposure in early life to high levels of PAHs may increase the risk of post menopausal breast cancer; however, other confounders related to geography cannot be ruled out.	Using extensive in-person interviews and self-administered questionnaires, participants provided information regarding medical history, diet, alcohol consumption, smoking history, lifetime passive smoke exposure, occupational history, and residential history. Residential histories were reported by the subject dating back to birth.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
10	**Perera F**, Rundle A. **Cancer Epidemiol Biomarkers Prev. 12(1):75-6. 2003.**	Letters to the editor	PAHs	/	/	/	significantly associated with breast cancer (odds ratio 2.56, 96% confidence interval 1.05–6.24; Carcinogenesis, Vol. 21, pp. 1281–1289, 2000)	/	/
11	**Kulldorff M,** Feuer EJ, Miller BA, et al. Am J Epidemiol. Jul 15;146(2):161-70.**1997.**	A geographic analysis. In the northeast **USA**	PAHs	Breast cancer clusters in the northeast United States	/	1988–1992	The New York City-Philadelphia metropolitan area had an excess of deaths among younger women at 6.9 percent compared with an excess of 7.4 percent in the older group.	The several known and hypothesized risk factors for which we could not adjust and that may explain the detected cluster are most notably age at first birth, age at menarche, age at menopause, breastfeeding, genetic mutations, and environmental factors.	The basic analysis is adjusted for age, with further analyses examining how the results are affected by incorporating race, urbanicity, and parity as confounding variables.
12	**Shen J**, Liao Y, Hopper JL, et al. Terry MB. Br J Cancer. 116(9):1229-1233. **2017.**	A prospective nested case-control study. New York **USA**	PAHs -albumin adducts	blood samples	After diagnosis	NR	Women with detectable levels of PAH had a twofold association with breast cancer risk (**OR = 2.04; 95% CI = 1.06–3.93**) relative to women with non-detectable levels.	These results support that family-based cohorts can be an efficient way to examine gene–environment interactions.	We measured plasma PAH-albumin adducts by competitive enzyme linked immune sorbent assay using monoclonal antibody 8E11 that recognises benzo(a)pyrene diolepoxide tetrols and related PAH metabolites.
13	**Lee KH**, Shu XO, Gao YT et al. Cancer Epidemiol Biomarkers Prev 19(3): 877–883. **2010.**	nested case-control study. In the Shanghai **China**	PAHs	Urinary	After diagnosis	1997 - 2000	No association was observed for PAH metabolites and the oxidative stress biomarkers of urinary malondialdehyde and 8-hydroxy-2'-deoxyguanosine and risk of breast cancer.	This nested case-control study provides no evidence of association between PAH exposure and oxidative stress and risk of breast cancer in Shanghai women.	Information on demographic characteristics, past medical history, lifestyles, history of menstruation, pregnancy history, eating and drinking habit, history of residence, family history.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
14	**Saieva C,** Peluso M, Masala G, et al. Breast Cancer Res Treat. 129(2):477-84. **2011.**	Prospective EPIC- study. **Italy**	PAHs	Blood	Before and after diagnosis	1993 - 1998	Overall, no significant difference in DNA adduct levels emerged between cases and their matched controls	Conditional regression analyses adjusted for selected potential confounders did not show any significant association between DNA adduct levels and BC risk. Thus, our results provide no evidence that bulky DNA adducts measured in peripheral leukocytes are associated with BC risk.	Age, BMI, smoking habits, education level, age at menarche, age at first delivery, and alcohol consumption

ANNEX 4 – References Polycyclic Aromatic Hydrocarbons (PAH) monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - PAHs	N°	REVIEWS - PAHs
1	**IARC Monographs. Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. Vol. 92: 1–853. 2010**	1	**Korsh J,** Shen A, Aliano K et al. **Breast Care. 10(5):316-8. 2015.**
2	**IARC Monographs. Bitumens and Bitumen Emissions, and Some N– and S-Heterocyclic Polycyclic Aromatic Hydrocarbons." Vol. 103. 2013**		

ANNEX 5 – References Perfluoroalkyl substances (PFASs, PFOS, PFOSA and PFOA) articles.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**Bonefeld-Jorgensen EC**, Long M, Bossi R, et al. **Environmental Health, 10:88, 2011.**	Case control study. **Greenland**	PFOA, PFOS	Blood Samples	After diagnosis	2000 - 2003	PFOS (ng/ml) OR = 1.03 (1.001; 1.07) *p = 0,05*	A case–control study of Greenland Inuit women found a positive statistically significant association between PFOS exposure and breast cancer but not statistically significant association between PFOA exposure and breast cancer.	Age, BMI, pregnancies, and cotinine;
2	**Ghisari M**, Eiberg H, Long M, et al. **Environ Health.** Mar 16;13(1):19. **2014.**	Case–control study. Greenlandic Inuit women **North Canadian**	PFOA, PFOS	Blood Samples	After diagnosis	2000 - 2003	Furthermore, an increased BC risk was observed for women with high serum levels of perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA)	The BRCA1 founder mutation and polymorphisms in CYP1A1 (Val) and CYP17 (A1) can increase the BC risk among Inuit women and the risk increases with higher serum levels of PFOS and PFOA	Age, BMI, smoking, menopause status, information on serum levels of PFOA / PFOS.
3	**Bonefeld-Jorgensen EC**, Long M, Fredslund SO, Bet al. **Cancer Causes Control. 25(11):1439-48. 2014.**	Case-cohort study. **Denmark.**	PFOA, PFOS	Blood Samples	15 years Before the diagnosis	1996-2002	PFOSA in the 5th quintile RR=2.40, 95% CI = 1.20, 4.83	The results of this study suggest an association between BC and the measured PFOSA	Age, BMI, smoking, menopause status, information on serum levels of PFOA / PFOS. Questionnaires on lifestyle and environmental exposure (including diet, height, weight, diseases in the family, smoking, and alcohol intake)

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
4	**Barry V**, Winquist A, Steenland K. **Environ Health Perspect. 121(11-12):1313-8. 2013;**	Cohort Analysis. Mid-Ohio Valley. **USA**	PFOA	Serum	After diagnosis	1992–2011	Breast Cancr HR = 0.93 (0.88, 0.99)	The results of this study suggest no association between BC and the measured PFOA	Proportional hazards modelling, using time varying cumulative exposure, adjusting for time-varying smoking, time-varying alcohol use, sex, education, 5-yr birth period;
5	**Vieira VM,** Hoffman K, Shin HM, et al. **EHP**,**Vol. 121, number 3. 2013;**	A Geographic Analysis. Parkersburg, West Virginia and Ohio. **USA**	PFOA	Serum	Before diagnosis	1996-2005	Estimated serum levels PFOA (µg/L) 10 yr before diagnosis (breast incidence) Low: 3.7–12.8 RR (95% CI) 0.9 (0.7–1.2) Medium: 12.9–30.7 RR (95% CI) 1.1 (0.8–1.5) High: 30.8–109 RR (95% CI) 0.7 (0.5–1.0) Very high: > 110 RR (95% CI) 1.4 (0.9–2.3) No evidence breast cancer	Our results suggest that higher PFOA serum levels may be associated with testicular, kidney, prostate, and ovarian cancers and non-Hodgkin lymphoma but no breast cancer	Age, BMI, smoking, information on serum levels of PFOA. Using geocoding, we were able to identify cases living within a contaminated water district area.
6	**Wielsoe M,** Kern P, Bonefeld-Jørgensen EC.**Environmental Health 16:56, 2017;**	Case-control study of Inuit women. **Greenland.**	PFAAS / PFOA	Serum	After diagnosis	2000-2003 and 2011-2014	Odds ratio of breast cancer risk associated with serum levels of PFAAs (ng/ml serum) PFOA OR (95% CI) 1.26 (1.01; 1.58) 1.00 (reference) 2nd Tertile 1.86 (0.80; 4.31) 3rd Tertile 2.64 (1.17; 5.97)	Significant, positive associations between breast cancer risk and PFOA and PFAAs were observed.	Information about age, body mass index (BMI), smoking status, menopause status, number of full term pregnancies, and history of breastfeeding was obtained from questionnaires.

ANNEX 5 – References Perfluoroalkyl substances (PFASs, PFOS, PFOSA and PFOA) monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - PFASs	N°	REVIEWS - PFASs
1	IARC Monographs. Vol. 110. 2010	1	**Steenland K,** Fletcher T, Savitz DA. **Environ Health Perspect. Aug;118(8):1100-8. 2010.**
		2	**Siddique S,** Kubwabo C, Harris SA. **Emerging Contaminants 2, 204-219. 2016.**

ANNEX 6 – References Triazine (atrazine, simazine, terbuthylazine and metabolites desethyl-atrazine, desethyl-terbuthylazine) articles.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**Kettles MA**, M K Kettles, S R Browning, et al. **105(11): 1222–1227. Environ Health Perspect.** 1997.	An ecologic study. Kentucky. **USA**	**Triazine**	water contaminant, corn planted, and pesticide use variables, counties were categorized by exposure status.	/	1991 - 1994	There is a statistically significant increase in breast cancer risk with medium and high levels of triazine exposure [OR = 1.14 p<0.0001 and OR = 1.2 p<0.0001].	The results suggest a relationship between female breast cancer incidence and triazine herbicide exposure.	Data on groundwater contamination. Data on surface water contamination. Survey of the amount of pesticide used by applicators in each county was examined
2	**McElroy JA**, Gangnon RE, Newcomb PA, et al. **J Expo Sci Environ Epidemiol,** 2007.	Case-control study. In rural areas in Wisconsin. **USA**	Atrazine	exposure estimation technique using a publicly available data set to examine atrazine exposure, in relation to breast cancer risk.	/	1988-2001	The possible risk for women exposed to levels of atrazine at or above statutory action levels of >3 ppb (OR 1.3, 95% CI 0.3–6.5) could not be ruled out due to small numbers in this category.	Resultsf rom this large population based study do not suggest an increased risk of breast cancer from adult exposure to atrazine in drinking water.	lifestyle risk factors: alcohol consumption and breast feeding habits; hormone replacement therapy use ; and finally, physical activity patterns.Analysiswasli mited to women who lived in rural areasof Wisconsin.
3	**Engel LS,** Hill DA, Hoppin JA, et al. Am J Epidemiol. 15;161(2):121-35. 2005.	Cohort study. In Iowa and North Carolina. **USA**	Atrazine and other pesticides	Pesticide exposure and Breast Cancer Risk	/	1993 - 1997	Breast cancer standardized incidence ratios were 0.87 (95% C.I.: 0.74, 1.02) for women who reported ever applying pesticides and 1.05 (95% C.I.: 0.89, 1.24) for women who reported never applying pesticides.	The authors found no clear association of breast cancer risk with farm size or washing of clothes worn during pesticide application, but risk was modestly elevated among women whose homes were closest to areas of pesticide application.	Questionnaires directed to the farmers elicited similar but more detailed information on lifetime pesticide use, including the duration and frequency of use of specific pesticides.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
4	**Rich CH**, Stump ML, Browning SR. **Arch Environ Contam Toxicol. Jan;42(1):127-36.2002.**	An ecologic study. Kentucky. **USA**	Atrazine	For this study, we assumed that the main route of atrazine exposure to the general population was through drinking water.	/	1993 - 1997	A null association was found for breast cancer across all exposure indices, both by county.	The results of this ecologic study did not generally support an association between exposure to the herbicide atrazine and breast cancer	An ecologic study was conducted using secondary data to derive measures of environmental exposure to atrazine in Kentucky
5	**Mills PK and Yang R. J Environ Health.68(6):15-22; 2006;**	Regression analysis California Latinas. **USA**	Triazine and other pesticides	Exposure at two classes of pesticides, organochlorines and triazine herbicides	/	1988-2000	No significant associations were found for the triazine herbicides atrazine and simazine.	Risk of breast cancer was inversely and significantly associated with fertility levels.	Age, socioeconomic status, and fertility rates.
6	**Hopenhayn-Rich C,** Stump ML, Browning SR.**Arch Environ Contam Toxicol 42:127–136. 2002;**	An ecologic study. Kentucky. **USA**	Atrazine	exposure to atrazine and the association between these measures and incidence of breast cancer	/	1993 - 1997	The ORs were 1.01 (95% CI 0.96, 1.05) and 0.98 (95% CI 0.93, 1.02) for the highest and next to- highest atrazine exposure groups, respectively.	A null association was found for breast cancer across all exposure indices.	The results of this ecologic study did not generally support an association between exposure to the herbicide atrazine and breast cancer
7	**Muir K,** Rattanamongkolgul S, Smallman-Raynor M, et al. **Public Health. 118(7):513-20. 2004;**	The ecological study. **UK**	Atrazine and other pesticides	breast cancer and environmental exposure to pesticides and the geographical distribution of pesticides	/	1989 - 1991	The ecologic studies do not support a causal relationship between breast cancer and atrazine.	Although pesticides do vary significantly in their spatial application, overall, consistent associations between breast cancer incidence rates and the pesticides applied were not found.	Information relating to all cases of female breast cancer (age .45 years at diagnosis), resident and diagnosed in Lincolnshire and Leicestershire in the period 1989–1991.
8	**Beane Freeman LE,** Rusiecki JA, Hoppin JA. **Environ Health Persp 119: 1253–1259. 2011;**	Cohort Study. Iowa and North Carolina. **USA**	Atrazine	Breast cancer incidence and its possible spatial association and exposure with pesticide application in two counties of England	/	1994 - 2007	there was no consistent evidence of an association between atrazine use and breast cancer.	In conclusion, despite high incidence of breast cancer in Lincolnshire and Leicestershire, the findings do not show localized spatial distribution of breast cancer incidence within these counties.	Information relating to all cases of female breast cancer, age (>45y), resident and diagnosed, pesticide application,

ANNEX 6 – References Triazine (atrazine, simazine, terbuthylazine and metabolites desethyl-atrazine, desethyl-terbuthylazine) monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - Triazine	N°	REVIEWS - Triazine
1	IARC Monographs. Vol. 73. 1999	1	Boffetta P, Eur J Cancer Prev, 2013;
		2	Simpkins JW, Swenberg JA, Weiss N, et al. Toxicol Sci. Oct;123(2):441-59. 2011.
		3	Sathiakumar, N., MacLennan, P. A., Mandel, J. Et al. Crit. Reviews Toxicol. 41(Suppl. 1), 1–34. 2011.

ANNEX 7 – References Cadmium (Cd) articles.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	Exposure category or level / Risk estimate (95% CI)	Relevant Comment	Methods and possible covariates
1	**Pan J**, Environ Geochem Health, **2010;**	Ecological study. **Europe**	Cadmium	European Geological Surveys geochemical baseline data to examine the distribution of cadmium (Cd) in Europe, with a particular reference to the international soil and water guideline values.	/	2010	The geochemical data used in this study were from the new Forum of European Geological Surveys (FOREGS), whereas the data on cancer incidence were from the International Agency for Research on Cancer (IARC) GLOBOCAN 2002 database.	In terms of human health impacts, food (up to several hundred lg/day) was found as the only major route of exposure to Cd for the non-smoking general population.	In this study, the incidence data on breast cancer for 26 European countries were used. In this study, the human health-related pathways of Cd were hence reviewed first
2	**Gallagher CM**, Research Paper, **2010;**	case-control study. **Long Island. USA**	urinary cadmium (UCd),	urine and blood samples	/	1999-2008	women in the highest quartile had increased risk for breast cancer (OR=2.69; 95% CI=1.07, 6.78) and US women in the two highest quartiles had increased risk (OR=2.50; 95% CI=1.11, 5.63 and OR=2.22; 95% CI=.89, 5.52, respectively).	Further research is warranted on the impact of environmental cadmium on breast cancer risk in specific populations and on identifying the underlying molecular mechanisms.	health questionnaire including family history of breast cancer, use of hormone therapy other than birth control pills, age at first live birth, and menopausal status, provided a blood sample.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	RR (95% CI)	Relevant Comment	Methods and possible covariates
3	**McElroy JA**, Shafer MM, Trentham-Dietz A, **J Natl Cancer Inst. 2006;**	case-control study. Wisconsin. **USA**	urinary cadmium (UCd),	urine samples	/	2004-2005	Women in the highest quartile of creatinine-adjusted cadmium level (≥ 0.58 µg/g) had twice the breast cancer risk of those in the lowest quartile (<0.26 µg/g; OR = 2.29, 95% CI = 1.3 to 4.2)	there was a statistically significant increase in risk with increasing cadmium level.	All participants were interviewed about physical activity, reproductive history, alcohol consumption, height and weight, use of oral contraceptives and hormone replacement therapy, personal and family medical history, demographic factors, a limited set of dietary components, and smoking history.
4	**Nagata C**, Nagao Y, Nakamura K. **Breast Cancer Res Treat. 2013;**	case-control study. **Japan**	urinary cadmium level	urine samples	after diagnosis (19 days)	2000-2002	Women in the highest tertile of the creatinine adjusted cadmium level ([2.620 lg/g) had significantly elevated OR of breast cancer relative to those in the lowest tertile (1.674 lg/g) after controlling for covariates OR = 6.05, (95 % CI 2.90, 12.62	We found that a higher urinary cadmium level was associated with an increased risk of breast cancer among Japanese women.	Age, BMI, menopausal status, and year of urine sampling, smoking and drinking habits, diet, physical activity, use of medication, medical history, and reproductive history

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	RR (95% CI)	Relevant Comment	Methods and possible covariates
5	**Strumylaite L**, Bogusevicius A, Abdrachmanovas O, et al. **Breast Cancer Res Treat. 2011;**	Study of (Cd) concentration in different biological media. **Lithuania**	cadmium (Cd)	breast adipose tissue, urine, and blood	after diagnosis	2007	Cd in urine was significantly higher in cancer patients than in controls (P < 0.001).	Cancer patients with positive estrogen receptors (ERs) had significantly greater concentration of breast tissue Cd compared to patients with negative ERs (P = 0.035).	Questionnaire demographic characteristics and smoking. The ER and progesterone-receptor (PR) levels were measured in the specimens of breast tissue from breast cancer
6	**Antila E,** Mussalo-Rauhamaa H, Kantola M, Atroshi F, Westermarck T. **Sci Total Environ 186:251-256. 1996;**	Case control study. **Finnish**	Association of cadmium with human breast cancer.	breast adipose tissue	after diagnosis	/	The cadmium concentrations found in the breast cancer patients did not differ statistically significantly from those of the healthy control	The postmenopausal breast cancer, decline in progesterone levels, and increased Cdtoxicity may be causally related.	In a questionnaire recording age, height, weight, smoking, fish-eating habits, parity, and previous breast-feeding.

ANNEX 7 – References Cadmium (Cd) monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - Cadmium	N°	REVIEWS - Cadmium
1	**IARC Monographs. Vol. 58. 1993**	1	**Jarup L, Berglund M, Elinder CG, et al. Scand J Work Environ Health 24(suppl 1):1–52. 1998.**

ANNEX 8 – References Trihalomethanes articles.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	RR (95% CI)	Relevant Comment	Methods and possible covariates
1	**Doyle TJ,** Zheng W, Cerhan JR,et al.. **Am J Public Health. Jul;87(7):1168-76.1997;**	A Prospective Cohort Study. Iowa. **USA**	trihalomethanes	An association between breast cancer risk and exposure of chlorination	/	1986 - 1993	An excess risk of breast cancer (RR = 1.33, 95% CI = 1.02, 1.74) was also observed in relation to surface water use.	Measurements of four trihalomethanes were performed on 252 municipal water supplies in Iowa.	Age, BMI, medical history, anthropometric data, and information concerning diet and risk factors for cancer, the types of drinking water.
2	**Koivusalo M,** Pukkala E, Vartiainen T, et al.**Cancer Cause Control. 8:192 – 200.1997;**	Cohort study **Finland.**	trihalomethanes	Drinking water chlorination and cancer-a historical cohort study in Finland.	/	1971 - 1993	RR an average exposure in a town using chlorinated surface level compared with those not exposed to chlorinated surface water. After adjustment for confounding, a statistically significant excess risk was observed for women in cancer of breast (RR = 1.11, CI = 1.01-1.22).	The increased risks were found primarily for women.	age, time period, urbanization, gender, and social-status
3	**Vinceti M,** Fantuzzi G, Monici L, et al. **Sci Total Environ. Sep 1; 330(1-3):47-53.2004;**	A retrospective cohort study in northern **Italy**	trihalomethanes	trihalomethane exposure through drinking water and cancer mortality in northern Italy	/	1987 - 1999	The risk of breast and ovarian cancer and of prostate cancer also tends to increase in subjects with higher socio–economic status, so the excess rates detected in our exposed cohort cannot be easily ascribed to life-style confounders.	There is epidemiologic evidence indicating that a higher socioeconomic status is directly associated with risk of breast cancer.	we also extracted all available information concerning the occupational status and educational attainment levels of this population at that date.

N°	Name of researchers / scientific journal	Place and study design	Exposure contaminants	Exposure assessment procedure	Time of sample collecting	Years of the Study interest	RR (95% CI)	Relevant Comment	Methods and possible covariates
4	**Ribera LF**, Lavedan EG, Aragonés N. et al. **Environment International 112;227–234.2018;**	multicase-control study **Spanish**	trihalomethanes, chloroform	Long-term exposure to trihalomethanes in drinking water and breast cancer in the Spanish multicase - control study on cancer	/	2008 - 2013	Adult-lifetime residential chloroform was associated with BC (adjusted OR =1.47; 95%CI =1.05, 2.06 for the highest (> 24 µg/L) vs. lowest (< 8 µg/L) quartile; p-trend =0.024)	At common levels in Europe, long-term residential total THMs were not related to female breast cancer. A moderate association with chloroform was suggested at the highest exposure category.	Age, residential history, water source in each residence (bottled, tap, other) and frequency and duration of bathing, educational level, occupational status, race, BMI, family history of BC, menopausal status, oral contraceptive use, age at menarche.
5	**Marcus PM,** Savitz DA, Millikan RC et al. **Epidemiology Vol. 9, No. 2, Mar., 1998;**	Ecological study. North Carolina. **USA**	trihalomethanes	Female breast cancer and exposure to trihalomethane levels in drinking water in North Carolina.	/	1995 - 1997	total trihalomethane category was not very different in black women (rate ratio = 1.2; 95% CI = 0.8-1.8) than in white women (rate ratio = 1.1; 95% CI = 0.9-1.3).	Total trihalomethane levels were not associated materially with breast cancer risk, adjusting for potential confounders.	ecologic measurements of age, income, education, urban status, and race as potential confounders.

ANNEX 8 – References Trihalomethanes monographs IARC, reviews, meta – analysis.

	MONOGRAPHS IARC - Trihalomethanes	N°	REVIEWS - Trihalomethanes
1	IARC Monographs. Vol. 101. 2013	1	Mohamadshafiee MR, Taghavi L. World Academy of Science, Engineering and Technology International Journal of Environmental and Ecological Engineering Vol:6, No:8, 2012.
2	IARC Monographs. Vol. 52. 1991		