



The quantification of social inequalities in the health burden of environmental stressors

Deliverable 3.2 of working package 3



Funded by
the European Union

BEST-COST is funded by the European Union's Horizon Europe programme under Grant Agreement No.101095408.



Deliverable Information	
Deliverable No. & Title	D3.2: the quantification of social inequalities in the health burden of environmental stressors
Work Package No.	3
Work Package Title	Social inequalities methodology
Lead Organisation	Sciensano, VITO
Main author(s)	Vanessa Gorasso (SCI), Frederik Priem (VITO), Gaëlle Mogin (SCI), Lien Poelmans (VITO), Jurgen Buekers (VITO)
Contributors	Alberto Castro (Swiss TPH), Axel Luyten (Swiss TPH), Carl Michael Baravelli (NIPH), Iracy Pimenta (UP), João Vasco Santos (UP), Sabrina Delaunay-Havard (SPF), Susanne Breitner (LMU)
External reviewer	University of Porto : João Vasco Santos, Andreia Novais, Iracy Pimenta
Nature	<i>Report</i>
Dissemination Level	<i>SEN (Sensitive)</i>
Deliverable Date	30/09/2024
Version Number	2



Contents

1	Disclaimer	9
2	Executive Summary	10
3	Introduction.....	11
3.1	Burden of environmental stressors	12
3.2	Social and environmental inequalities	13
3.3	Objectives and structure of the report	13
4	Data availability	16
4.1	Population and demography.....	16
4.1.1	EU availability	16
4.1.2	Case country availability	17
4.2	Environmental stressors	19
4.2.1	EU availability	19
4.2.2	Case country availability	20
4.3	Health outcomes	23
4.3.1.1	EU availability.....	25
4.3.1.2	Case country availability	25
4.4	Deprivation	31
4.4.1	EU and case country availability	32
4.5	Summary and highest common resolution.....	35
5	Methodological framework for data mapping	36
5.1	Outline of the proposed framework	36
5.2	General mapping protocol.....	37
5.2.1	Software and data	37
5.2.2	Support graphs.....	37
5.2.3	Data processing	37
5.2.4	Map design guidelines.....	38
5.2.5	Map extent and Coordinate Reference Systems	39
5.3	Mapping approach per data availability scenario	39
5.3.1	Scenario 1 – Mapping a single high-resolution indicator.....	39
5.3.1.1	Selected example data.....	39
5.3.1.2	Mapping	39



5.3.1.3	Support graph	42
5.3.2	Scenario 2 – Mapping a single low resolution indicator	44
5.3.2.1	Selected example data.....	44
5.3.2.2	Mapping	45
5.3.2.3	Support graph	48
5.3.3	Scenario 3 – Mapping two equal resolution indicators	50
5.3.3.1	Selected example data.....	50
5.3.3.2	Mapping	51
5.3.3.3	Support graph	54
5.3.4	Scenario 4 – Mapping a higher and lower resolution indicator	56
5.3.4.1	Selected example data.....	56
5.3.4.2	Mapping	57
5.3.4.3	Support graph	58
5.3.5	Scenario 5 – Mapping three equal resolution indicators	60
5.3.5.1	Selected example data.....	60
5.3.5.2	Mapping	60
5.3.5.3	Support graph	64
5.4	Concluding remarks on the mapping	65
6	Conclusion and recommendations	66
7	References	68
	Appendix	71



List of Tables and Figures

Figure 3-1: visual representation of WP3 tasks	11
Figure 3-2: Leading level 3 risk factors by attributable DALYs per 100,000 (1990 and 2021).	12
Table 3-1 : Geographical levels corresponding to the NUTS and LAU of the case countries, and the number of units for each level.....	14
Table 4-1 : Geographical area characteristics, population demographics and surface area details for the five case countries and the EU.	18
Table 4-2: Air pollutants (PM2.5, O3 and NO2) and noise pollution (from road traffic, railway and aircraft) for the five case country. For each pollutant, metric and unit, the geographical area or resolution, the geographical coverage, the reference period and the data type are displayed.....	21
Table 4-2: continuation (Norway, Portugal and EU)	22
Table 4-3: Definition of diseases using ICD codes	24
Table 4-4: Data availability of disease prevalence and/or incidence. The disease level, smallest geographical area and last available reference period are displayed by country.	27
Table 4-4 (continuation – France, Norway and Portugal)	28
Table 4-5 : Data availability of disease mortality (EU, Belgium and Estonia). The disease level, smallest geographical area, and last available reference period are displayed.....	29
Table 4-5 (continuation) : France, Norway and Portugal	30
Table 4-6: Data availability for the five selected deprivation indicators in the five countries concerned (Belgium, Estonia, France, Norway, Portugal) and in the European Union. For each country, the indicator is described and the lowest geographical level at which it is available is indicated.	33
Table 4- 6: continuation	34
Table 4-7: Summary table for pollutant exposure and disease outcome considering the highest geographical resolution available for each country and the EU.	35
Figure 5-1. Choropleth map with a bivariate colour scheme showing 2021 NO ₂ air pollution on 1km ² resolution in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted here	40
Figure 5-2. Choropleth map with a bivariate colour scheme showing 2021 NO ₂ air pollution exposure on NUTS3-level in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted here	42
Figure 5-3. Colour coded histogram showing the quantitative distribution of the 2021 NO ₂ air pollution exposure on NUTS3-level mapped in Figure 5-2 . A high-resolution version of this figure can be consulted here	43



- Figure 5-4.** Colour coded horizontal bar chart showing 2021 national-level (NUTS0) NO₂ air pollution exposure in the EU27, its Candidate Countries and the UK. The bar colour coding matches that map layout of **Figure 5-2**. A high-resolution version of this figure can be consulted [here](#). 44
- Figure 5-5.** Proportionate symbol map, using circle diagrams, showing 2021 age-weighted mortality due to malignant neoplasms on NUTS1-level in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#). 45
- Figure 5-6.** Pie chart map showing 2021 total age-weighted mortality and age-weighted mortality by cause of death on NUTS1-level in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#). 46
- Figure 5-7.** Line pattern map showing 2021 age-weighted mortality due to malignant neoplasms on NUTS1-level in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#). 47
- Figure 5-8.** Histograms showing the quantitative distribution of 2021 age-weighted mortality due to all causes of death, malignant neoplasms, diseases of the circulatory system and other causes of death on NUTS1-level in the EU27 and its Candidate Countries. The colouring of the bars matches the respective class colours used in **Figure 5-6**. A high-resolution version of this figure can be consulted [here](#). 49
- Figure 5-9.** Stacked horizontal bar chart showing the 2021 national-level (NUTS0) overall age-weighted mortality and by cause of death in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#). 50
- Figure 5-10.** Choropleth with a unipolar colour scheme showing 2021 NO₂ air pollution exposure on NUTS2-level in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#). 52
- Figure 5-11.** Choropleth map with a unipolar colour scheme showing 2021 age-weighted mortality due to diseases of the circulatory system on NUTS2-level in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#). 53
- Figure 5-12.** Bivariate choropleth map simultaneously showing 2021 NO₂ air pollution exposure and age-weighted mortality due to diseases of the circulatory system on NUTS2-level in the EU27, its Candidate Countries and the UK. The bivariate colour scheme is obtained by overlaying the unipolar colour schemes used in **Figure 5-10** and **Figure 5-11**. A high-resolution version of this figure can be consulted [here](#). 54
- Figure 5-13.** Colour coded scatterplot with lateral histograms showing the quantitative (co-) distributions of 2021 NO₂ air pollution exposure and age-weighted mortality due to diseases of the circulatory system on NUTS2-level in the EU27 (excluding Turkey) and its Candidate Countries. The colour coding of the scatterplot matches the bivariate colour scheme used in **Figure 5-12**. A high-resolution version of this figure can be consulted [here](#). 55
- Figure 5-14.** Mirrored horizontal bar chart showing 2021 national-level (NUTS0) NO₂ air pollution exposure, on the left, and age-weighted mortality due to diseases of the circulatory system, on the right, on NUTS2-level in the EU27 and its Candidate Countries (excluding Turkey). A high-resolution version of this figure can be consulted [here](#). 56



Figure 5-15. Combined line pattern map and choropleth map with a unipolar colour scheme, respectively showing 2021 age-weighted mortality due to diseases of the circulatory system, on NUTS1-level, and noise pollution exposure obtained with the Quietness Suitability Index, on NUTS3-level, in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#). 58

Figure 5-16. Colour coded and hatched scatterplot with lateral histograms showing the quantitative (co-)distribution of 2021 age-weighted mortality due to diseases of the circulatory system, on NUTS1-level, and 2016 population-weighted complement of the Quietness Suitability Index, expressing noise pollution exposure on NUTS3-level, in the EU27, its Candidate Countries (excluding Turkey) and the UK. The colour coding and hatch patterns of the different classes match those used in the map layout of **Figure 5-15**. A high-resolution version of this figure can be consulted [here](#). 59

Figure 5-17. Choropleth map showing 2021 Severe Material and Social Deprivation on NUTS2-level in the EU27, its Candidate Countries and the UK. The four SMSD quantiles classes mapped here are re-used to define the focus areas of the subsequent maps **Figure 5-18** - **Figure 5-21**. A high-resolution version of this figure can be consulted [here](#). 61

Figure 5-18. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a very low score ($\text{SMSD} < 2\%$) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#). 62

Figure 5-19. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a low score ($2\% \leq \text{SMSD} < 5\%$) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#). 62

Figure 5-20. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a moderate score ($5\% \leq \text{SMSD} < 10\%$) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#). 63

Figure 5-21. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a high score ($\text{SMSD} \geq 10\%$) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#). 63

Figure 5-22. Colour coded scatterplot with lateral histograms, shown and described earlier in **Figure 5-13**, now having plot symbols whose shapes and colours reflect the varying degrees of Severe Material and Social Deprivation of the corresponding NUTS2 regions. A high-resolution version of this figure can be consulted [here](#). 64



Table of abbreviations

BEST-COST	Burden of disease based methods for estimating the socioeconomic cost of environmental stressors
DALY	Disability-Adjusted Life Years
DCS	Disease of the Circulatory System
dB	Decibel
EEA	European Environment Agency
EU	European Union
EUROSTAT	Statistical office of the European Union
GBD	Global Burden of Disease
GIS	Geographic Information Systems
ICD	International Classification of Diseases
IRIS	Ilots Regroupés pour l'Information Statistique
LAU	Local Administrative Units
Lden	Day–evening–night noise level
Lnight	Night-time noise (Lnight)
MDI	Multiple Deprivation Index
NO ₂	Nitrogen dioxide
NUTS	Nomenclature of territorial units for statistics
O ₃	Ozone
PM _{2.5}	Particles with a diameter of 2.5 micrometres or less
SMSD	Severe Material and Social Deprivation
USA	United States of America
WHO	World Health Organization



1 Disclaimer

The opinions in this report reflect the opinions of the authors and not the opinions of the European Commission. The European Union is not liable for any use that may be made of the information contained in this document.

All intellectual property rights are owned by the BEST-COST consortium members and are protected by the applicable laws. Except where otherwise specified, all document contents are: “© BEST-COST project - All rights reserved”. Reproduction is not authorised without prior written agreement.

The commercial use of any information contained in this document may require a license from the owner of that information.

All BEST-COST consortium members are also committed to publish accurate and up to date information and take the greatest care to do so. However, the BEST-COST consortium members cannot accept liability for any inaccuracies or omissions, nor do they accept liability for any direct, indirect, special, consequential or other losses or damages of any kind arising out of the use of this information.



2 Executive Summary

The overall objective of BEST-COST is to improve the methodology for assessing the socioeconomic cost of environmental stressors. Environmental stressors disproportionately affect socially disadvantaged groups, and, therefore, contribute to socioeconomic inequalities in the burden of disease. Despite the increased attention to inequalities in health and the environment, there is currently no systematic monitoring of environmental health inequalities. To overcome this, BEST-COST Work Package 3 aimed to develop and implement an innovative and coherent methodological framework for assessing socioeconomic inequalities in the health impact of environmental stressors with a focus on air pollution and traffic-related noise. This report firstly aims to assess the availability of (1) exposure to environmental stressors, (2) health outcomes in terms of mortality and morbidity and (3) socioeconomic deprivation data in Europe. The BEST-COST project also targets the highest geographical resolution possible to capture local variations of environmental stressors, health outcomes and deprivation. A key challenge identified in this task was the variability in data availability and geographic resolution across countries, which hindered the initial goal of producing high-resolution burden estimates for the entire European Union. To address this, two recommendations were made: (1) researchers need to decide whether to focus on specific outcomes (e.g. mortality) or limit the analysis to one country in order to achieve a higher resolution, and (2) use low-resolution data when comparing multiple countries or regions to avoid information loss. For this project, it seems feasible to calculate the socioeconomic burden of environmental stressors at the NUTS 2 level. The second part of the report addresses the cartographic mapping of the targeted statistics. The key objective here is to develop a methodological framework that can be applied generically on various indicators or combinations to help produce maps that provide spatial insight into these data and their interrelations. A total of five data availability scenarios are included in this framework to help researchers and analysts produce clear and meaningful maps of the socio-economic burden of environmental stressors



3 Introduction

The overall objective of BEST-COST is to **improve the methodology for the assessment of the socioeconomic cost of environmental stressors** to i) enhance regular usage of economic and health modelling in policy impact assessments and policy evaluation by the European Union (EU) and national public authorities, and ii) promote harmonised and consensual population health, quality of life and economic metrics for integrative socioeconomic assessments of environmental pollution in Europe and health impact and cost-benefit assessments of related policies. The BEST-COST project comprises a total of nine work packages involving 18 organisations from 10 European countries (Belgium, Switzerland, Germany, Denmark, Estonia, Finland, France, The Netherlands, Norway, Portugal) and the USA.

Within BEST-COST, working package 3 (WP3) aims to 1) develop a coherent methodological framework for assessing the extent of social inequalities in different EU countries; 2) provide a methodological framework to quantify the burden and cost associated with exposure to environmental stressors; 3) quantify the contribution of social inequalities in the burden of environmental stressors based on a novel index of multiple deprivation. **Figure 3-1** visually depicts the combination of the three cornerstones of WP3 (social deprivation, environmental stressors and socioeconomic cost) and their combination in the tasks.

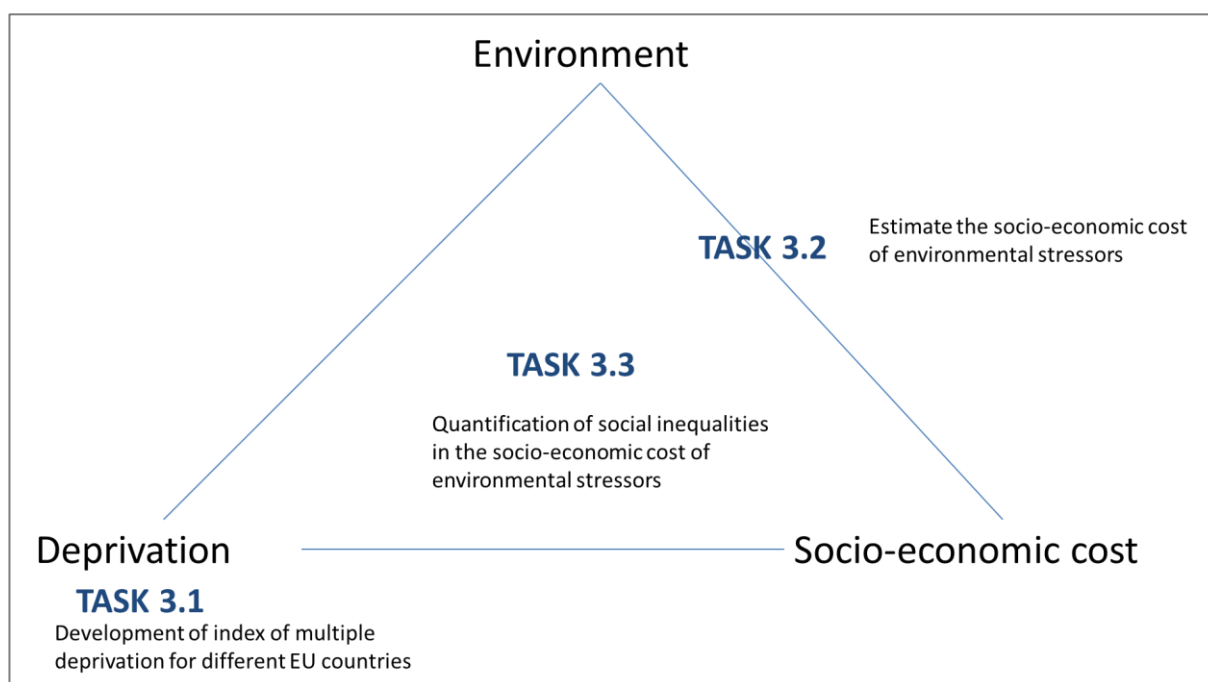


Figure 3-1: visual representation of WP3 tasks



3.1 Burden of environmental stressors

The GBD defines **risk factors** as “an attribute or exposure which is causally associated with an increased probability of a disease or injury” [1]. Some examples of important behavioural and metabolic risk factors include tobacco consumption, high alcohol consumption, high blood pressure, high fasting plasma glucose, a high body mass index and a diet low in fruits and vegetables [2].

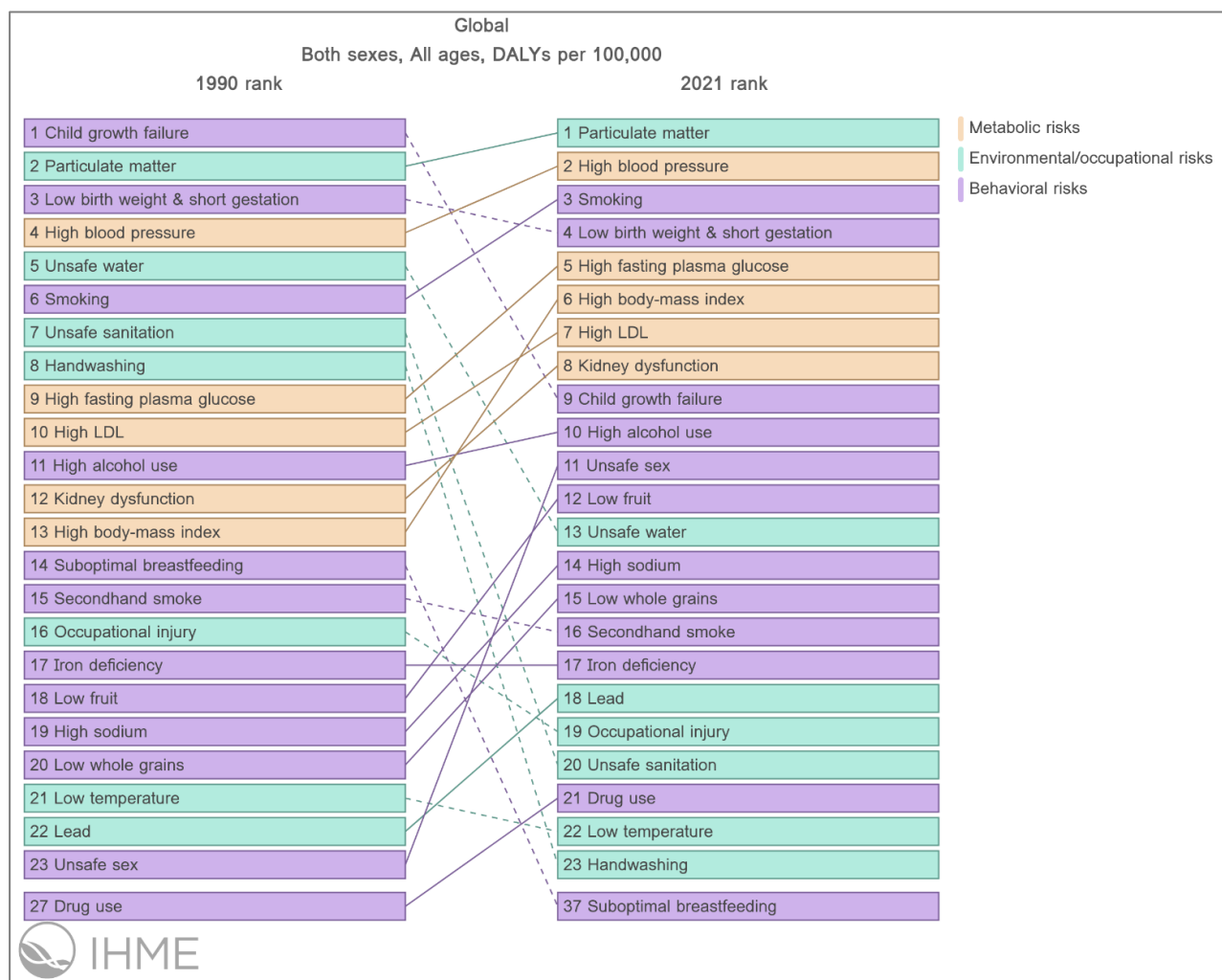


Figure 3-2: Leading level 3 risk factors by attributable DALYs per 100,000 (1990 and 2021).

Many diseases and deaths can also be attributed to **environmental risk factors**, including air pollution, noise, land-use patterns, the working environment and climate change [3]. According to the 2021 Global Burden of Disease (GBD) [4] study estimates, particulate matter air pollution was the leading contributor to the global disease burden in 2021 (see **Figure 3-2**), as measured by age-adjusted disability-adjusted life years (DALYs), which accounts for both fatal and non-fatal causes, and seventh contributor to the European region, including the leading risk factor for DALYs among all environmental and occupational risks. Particulate matter air pollution was estimated to contribute to 8% of total DALYs and 7.83 million deaths worldwide in 2019 [4], [5].



Individuals can further be exposed to environmental noise from traffic or street works, which may be an additional significant source of disease burden not yet included in the GBD study. Moreover, exposure differences to environmental stressors such as chemicals, heat waves or radiations in our daily lives may impact health across the population.

3.2 Social and environmental inequalities

Research suggests that exposure to environmental risk factors may vary among different socioeconomic groups, contributing to health inequalities [6]. It has been demonstrated that in Europe, vulnerable populations such as children, the elderly, people with lower socioeconomic status and those in poor health are more likely to experience adverse health outcomes from environmental exposures [7], [8]. It is, therefore, important to quantify social determinants of health inequalities in order to identify population groups that should be specifically targeted by health policies. Measuring **deprivation**, defined as “*a state of observable and demonstrable disadvantage relative to the local community or the wider society to which an individual, family or group belongs*”, is an effective way to assess social inequalities [9].

In line with the attention drawn to social inequalities in health, increased attention has been drawn to **environmental inequalities** – i.e., differences in the levels of environmental exposure between groups of people according to their socioeconomic status. Quite often, this is referred to as environmental justice [10]. As with social inequalities in health, environmental inequalities are widespread and persistent, as was recently inventoried by the World Health Organization (WHO) Regional Office for Europe [11].

Despite the increased attention to social inequalities in health and the environment, there is no systematic monitoring of environmental health inequalities, i.e., the contribution of social inequalities to the health impact of environmental stressors. To overcome this, the BEST-COST project aimed to develop and implement an innovative framework for assessing social inequalities in the health impact of environmental stressors using a novel multiple deprivation index (MDI). The latter was developed under the previous task within WP3 (for more information, see report 3.1).

3.3 Objectives and structure of the report

This report firstly aims to assess the data availability of (1) exposure to environmental stressors, (2) health outcomes in terms of mortality and morbidity and (3) socioeconomic deprivation data in Europe. These objectives are elaborated in the first chapter of this report. A comprehensive set of relevant indicators was selected for each of the three dimensions (environmental exposures, health outcomes, socioeconomic deprivation), for which data availability was explored in national-scale datasets for the five selected case countries: Belgium, Estonia, France, Norway and Portugal. In addition, indicators available through the EU institutions, notably the European Union statistical office (EUROSTAT) and the European Environment Agency (EEA), were screened. Particular attention was paid to the spatial, temporal, and thematic resolutions of the targeted statistics. In the final section of this chapter, the highest common resolutions among EU- and case country-scale was discussed,



considering the importance for BEST-COST of being able to combine the different dimensions at a high geographical resolution.

The BEST-COST project targets the **highest geographical resolution possible** to be able to capture local variations of deprivation, pollutant exposures and disease occurrence. Differences in population characteristics can help identify relationships between environmental exposures and health data. To reference countries' regions for statistical purposes, the EU has developed a classification known as NUTS (Nomenclature of territorial units for statistics) [12] and LAU (Local Administrative Units) [13].

Text box 1 - Geographical scale refers to the level of spatial detail or extent at which geographic data is analyzed or represented. It ranges from broad, large-scale perspectives that cover extensive areas (low resolution) to more localized, small-scale views that provide greater detail (high resolution). A higher resolution captures finer geographic features, while a lower resolution offers a more general, less detailed overview of a larger area.

In the early 1970s, Eurostat introduced the **NUTS** classification as a unified and consistent framework for dividing the EU's territory to facilitate the production of regional statistics. NUTS divides each EU country into 3 levels (from largest to smallest):

- NUTS 1: major socioeconomic regions
- NUTS 2: basic regions (for regional policies)
- NUTS 3: small regions (for specific diagnoses)

LAUs (LAU 1 and LAU 2 levels) are the building blocks of NUTS and comprise the EU's municipalities and communes. They are sometimes defined as level NUTS 4 and NUTS 5. However, it should be noted that since 2017, there is only one LAU level, but this report sometimes uses the old designations depending on the date of the data. **Table 3-1** highlights the geographical levels corresponding to the NUTS (2024) and LAU (2022) of the case countries, and the number of units for each level:

More information on the establishment of a common classification of territorial units for statistics (NUTS) by the EU can be found [here](#).

Table 3-1: geographical levels corresponding to the NUTS and LAU of the case countries, and the number of units for each level. Source : <https://ec.europa.eu/eurostat/web/nuts/correspondence-tables>;

*Geographical level names are translated into English ; ** AR = Autonomous Regions

	NUTS 1		NUTS 2		NUTS 3		LAU	
	Name*	Units	Name*	Units	Name*	Units	Name*	Units
Belgium	Regions	3	Provinces	11	Districts	44	Municipalities	581
Estonia	-	1	-	1	County groups	5	Municipalities	79
France	Regions	14	Regions + collectivities	27	Departments	101	Municipalities	34,956
Norway	-	1	Regions	7	Counties	17	Municipalities	378
Portugal	Mainland + AR**	3	Groups of intermunicipal entities + AR	9	Intermunicipal entities + AR	26	Parishes	3,092



The second chapter of the report addresses the cartographic mapping, using Geographic Information Systems (GIS), of the targeted statistics, which all have a pronounced spatiotemporal component. The key objective here is to develop a methodological framework that can be applied generically to various indicators or combinations thereof to help produce maps that provide spatial insight into these data and their interrelations. First, we outlined a total of five mapping scenarios with matching mapping protocols. These five mapping scenarios were identified along the lines of the spatial resolution on which the mapped data are made available and the number of indicators that are to be studied simultaneously within a single map or map series. The five scenarios form the backbone of the framework and address various mapping requirements that could commonly be encountered when working on the three interlinked core topics of the BEST-COST project, i.e., environmental stressors, health outcomes and socioeconomic deprivation. Each of the five mapping scenarios is elaborated with a concise explanation of the suitable mapping approach(es) and illustrated with some example maps and support graphs. The code, data organisation and other technical details of the developed mapping workflows will be disclosed via the BEST-COST protocols and GitHub channel.

In the third and final chapter, we formulated the concluding remarks of this report and put forward several recommendations to assist the other work package consortia of BEST-COST in their further endeavours.



4 Data availability

The case countries, air and noise pollutants and diseases of interest were previously defined based on the work of the BEST-COST project, namely WP1 and WP3. Based on their findings, we assembled extraction sheets to collect the meta-data of the selected indicators. Particular attention was given to identifying the highest geographical resolution by which the socioeconomic cost of environmental stressors can be quantified in each case study country.

Text box 2 - High geographical resolution. While there is no convention as to what constitutes a high or low spatial resolution, it was decided in the context of the BEST-COST project, for reference and considering the typical resolutions of the datasets treated in this report, to label NUTS1- as lower resolution and NUTS2+ as higher resolution. In any case, this definition is only indicative. Spatial data distributed in raster format with pixel sizes of 1km or smaller are also considered having a higher resolution for the purpose of this report.

Screening of data availability was performed by national partners within each of the case study countries in two rounds. One for population information, environmental stressors and health outcomes, and the second one for deprivation indicators. They were provided with the extraction sheet (example for Belgium can be found in **Appendix 2**) and a manual for filling in the extraction (**Appendix 1**).

4.1 Population and demography

General information regarding the country and the geographical units commonly used to divide its territory (regions, counties, municipalities, other smaller statistical areas or census tracts) were extracted (for example, see **Appendix 2.1**). For each geographical area, we indicated the corresponding European NUTS [12] or LAU [13] level, if applicable. In the case of country-specific geographical areas (e.g. IRIS in France), partners provided a short description of the geographical area. For each administrative division, the number of entities, as well as the average, and the minimum and maximum range of the surface area and of the population were extracted.

We also extracted information on the availability of population grids (maps that show population numbers or densities in regular grid cells) for the countries of interest. These might be useful for converting pollution levels in a spatial format into the area-level exposure needed for disease burden calculations (e.g. a population-weighted average concentration of pollutants).

4.1.1 EU availability

The EU data comes from Eurostat. Eurostat covers 42 countries, 92 units at NUTS1 level, 242 units at NUTS2 level, and 1166 units at NUTS3 level (Table 4-1). In 2023, the average population was close to 14 million at the national level, and 380,941 at the NUTS3 level. The average surface area is also much smaller at the NUTS3 level (3,877 km²) than at the average national level (160,656 km²).



4.1.2 Case country availability

All countries are divided into at least three geographic units, including regions, provinces, areas, municipalities, etc (**Table 4-1**). The smallest geographical division of each country is the LAU2 level, corresponding to municipalities or equivalent [13]. It is important to note that for some countries, even if data is technically available, acquiring it at a small scale could be practically difficult due to costs. Some countries have a higher geographical resolution, such as Belgium with its statistical sectors (19,795 units) and France with its IRIS (Ilots Regroupés pour l'Information Statistique - 48,577 units). The average size of the various geographical units varies greatly from one country to another. For example, the average size of a municipality is 2.9 km² in Belgium, 5,502 km² in Estonia, 15 km² in France, 855 km² in Norway and 1,048 km² in Portugal. The average population of municipalities also varies greatly from country to country, from an average of 1,901 inhabitants in France to 54,416 in Portugal

Table 4-1 : Geographical area characteristics, population demographics and surface area details for the five case study countries and the EU in 2023.

Geographical area characteristics				Population			Surface area (km²)		
Country	Geographical area	Corresponding NUTS or LAU	Number of units	Mean	Min	Max	Mean	Min	Max
Belgium	Regions	NUTS 1	3	3,872,541	1,228,655	6,709,787	10,222	162	16,906
	Provinces	NUTS 2	11	1,038,896	294,400	1,890,627	2,788	162	4,461
	Districts	NUTS 3/ LAU 1	44	268,339	50,896	1,228,655	697	162	1,597
	Municipalities	LAU 2	581	20,254	77	543,165	2.9	1.2	215
	Statistical sectors*	/	19,795	569	0	8,569	1.3	0.01	44.8
Estonia	Country	NUTS 2	1	-	-	1,331,796	-	-	45,339
	Areas**	NUTS 3	5	266,359	121,931	614,561	8,686	3,364	15,533
	Counties	LAU 1	15	61,313	8,330	372,243	2,898	1,032	4809
	Municipalities	LAU 2	79	16,858	89	437,811	5,502	3.8	2,717
France	Regions	NUTS 1	13	5,064,218	351,255	12,358,932	42,424	8,722	85,184
	Departments	NUTS 3	96	685,779	76,648	2,606,646	5,720	105	10,378
	Municipalities	LAU 2	34,816	1,901	0	498,596	15.8	0.03	757.08
	IRIS***	/	48,577	1,363	0	15,915	11.3	0.01	456
Norway	Country	NUTS 1	1	5,435,536	5,435,536	5,435,536	304,045	304,045	304,045
	Administrative Regions	NUTS 2	6	914,831	373,628	2,001,278	-	-	-
	Counties	NUTS 3	11	498,999	241,084	1,292,241	-	-	-
	Economic regions	NUTS 4 / LAU 1	85	64,576	9,963	709,037	-	-	-
	Municipalities	NUTS 5 / LAU 2	356	15,373	208	709,037	855	6	8,968
Portugal	Regions	NUTS 1	3	1,181,928	239,368	9,951,898	1,044	801	89,103
	Administrative Regions	NUTS 2	7	855,785	239,368	3,620,740	783	801	31,605
	Intermunicipal Entities	NUTS 3	25	80,584	423	2,891,658	896	801	8,543
	Municipality/Counties	LAU 1	308	51,416	423	546,923	1,048	8	1,721
	Parishes	LAU 2	4,257	2,478.9	31	66,250	29.8	0.2	888.35
EU	Country	/	42	13,803,291	33,812	85,279,553	160,656	160	783,562
	NUTS 1	NUTS 1	92	4,648,256	1,746	17,924,591	49,399	160	336,859
	NUTS 2	NUTS 2	242	1,783,633	1,746	12,329,432	17,894	13	227,150
	NUTS 3	NUTS 3	1,166	380,941	1,746	6,726,640	3,877	13	105,205

*census tract , **consists of different counties and *** = Ilots Regroupés pour l'Information Statistique



4.2 Environmental stressors

Exposure to fine particulate matter (PM_{2.5}), ozone (O₃) and nitrogen dioxide (NO₂) as well as exposure to transportation noise from roads, railways and aircraft are the focus of BEST-COST. For each pollutant dispersion data source, each case country was asked to indicate:

- The metric and unit that expresses the level of pollution (e.g., number or density);
- The geographical area for which exposure is available directly, or the resolution of the raster (e.g., the cell size of the PM_{2.5} or population grid) or vector map (e.g., the width of the noise contours) that will form the basis of the exposure assessment. It was possible to indicate up to three sources, the smallest areas or highest resolutions being the most important for the study;
- The completeness of the exposure numbers (e.g., nationwide, regional...) that is covered by the pollution map;
- The reference period (e.g., year), or if the data was available for multiple periods, the most recent year;
- The type of the data (example: meteorological station, administrative data, model output, satellite image);
- The accessibility of the data, by choosing one of the following options:
 - The data is owned by the institution of the user and can be used freely;
 - The data is owned by another institution but is available as open access;
 - The data is owned by another institution and can be used given authorisation of the owner (e.g., by signing a contract);
 - The data is owned by another institution and cannot be used.
- The source of the data (institution in charge) and a weblink to the data in case of open access.
- Whether exposure numbers are stratified by age and/or sex (not applicable for raster or vector data) and if so, specify the strata (e.g., definition of age groups);
- Whether socioeconomic indicators accompany the exposure numbers and if so, specify which ones (not applicable for raster or vector data).
- Additional information if needed (e.g. if data is accessible on a smaller scale but requires authorization).

See **Appendix 2.2** for an example.

4.2.1 EU availability

All data concerning concentrations of air and noise pollutants on a European scale come from the [EEA website](#). Atmospheric pollutant concentrations of PM_{2.5} (µg/m³), O₃ (ppb) and NO₂ (µg/m³) are available for 38 countries on a very small geographical scale (1km*1km) (Table 4-2). The reference period is 2021, and the data are based on a gridded-model output (regression-interpolation-merging mapping). For road traffic, railways and aircraft noise, emissions are calculated using Lden (day–evening–night noise level) or Lnight (night-time noise), both in dB. Data are available for 38 countries at agglomeration, regional or national





level, depending on the country. The reference period is 2016, and the data are based on an exposure number per 5dB band.

4.2.2 Case country availability

Most countries use the annual average concentration to estimate air pollutant concentration levels, and Lden or Lnight to measure noise exposure (Table 4-2). Geographical resolution for air pollutants varies from country to country, ranging from 10m*10m in Belgium to municipality level in Norway and Portugal. For noise exposure, the geographical resolution is smaller, generally representing exposure at major traffic arteries and cities or at regional level.



Table 4-2: Air pollutants (PM_{2.5}, O₃ and NO₂) and noise pollution (from road traffic, railway and aircraft) for the five case study country. For each pollutant, metric and unit, the geographical area or resolution, the geographical coverage, the reference period and the data type are displayed.

Country	Exposure	Metric and unit	Geographical area/resolution	Geographical coverage	Reference period	Data type
Belgium	PM _{2.5}	Annual mean concentration in µg/m ³	10m*10m	Belgium	2021	Gridded model output (ATMO-Street)
	O ₃	Annual mean of daily maximum 8-hour mean concentration in µg/m ³	10m*10m	Belgium	2021	Gridded model output (RIO-IFDM)
	NO ₂	Annual mean concentration in µg/m ³	10m*10m	Belgium	2021	Gridded model output (ATMO-Street)
	Road traffic	Lden or Lnight in dB	Regions	Regions	2016	Exposure number per 5 dB band
	Railway	Lden or Lnight in dB	Regions	Regions	2016	Exposure number per 5 dB band
	Aircraft	Lden or Lnight in dB	Regions	Regions	2016	Exposure number per 5 dB band
Estonia	PM _{2.5}	Annual mean concentration in µg/m ³	Kohtla-Järve, Tallinn and Tartu	Kohtla-Järve, Tallinn and Tartu	2022	Graph, table
	O ₃	Annual mean of daily maximum 8-hour mean concentration in µg/m ³	Kohtla-Järve, Tallinn and Tartu	Kohtla-Järve, Tallinn and Tartu	2022	Graph, table
	NO ₂	Annual mean concentration in µg/m ³	Kohtla-Järve, Tallinn and Tartu	Kohtla-Järve, Tallinn and Tartu	2022	Graph, table, dispersion model
	Road traffic	Lden or Lnight in dB	Tallinn and Tartu Region	Tallinn and Tartu Region	2022	Spatial data, modelled
	Railway	Lden or Lnight in dB	Tallinn and Tartu Region	Tallinn and Tartu Region	2022	Spatial data, modelled
	Aircraft	Lden or Lnight in dB	Tallinn and Tartu Region	Tallinn and Tartu Region	2022	Spatial data, modelled
France	PM _{2.5}	Hourly/daily/monthly/yearly concentrations (µg/m ³)	4km*4km	Metropolitan France	2000-2024	Modelled concentrations
	O ₃	Hourly/daily/monthly/yearly concentrations (µg/m ³)	4km*4km	Metropolitan France	2000-2024	Modelled concentrations
	NO ₂	Hourly/daily/monthly/yearly concentrations (µg/m ³)	4km*4km	Metropolitan France	2000-2024	Modelled concentrations
	Road traffic	Lden or Lnight in dB	Specific zones around source of noise	Metropolitan France	2017	Modelled noise pollution from the specified source
	Railway	Lden or Lnight in dB	Specific zones around source of noise	Metropolitan France	2017	Modelled noise pollution from the specified source
	Aircraft	Lden or Lnight in dB	Specific zones around source of noise	Metropolitan France	2017	Modelled noise pollution from the specified source



Table 4-2: continuation (Norway, Portugal and EU)

Country	Exposure	Metric and unit	Geographical area/resolution	Geographical coverage	Reference period	Data type
Norway	PM _{2.5}	Annual mean concentration in µg/m ³	Country/County/Municipality	Norway	2016-2018	Exposure number per 1 ug/m3 category
	O ₃	Annual mean concentration in µg/m ³	Country/County/Municipality	Norway	2016-2018	Exposure number per 1 ug/m3 category
	NO ₂	Annual mean concentration in µg/m ³	Country/County/Municipality	Norway	2016-2018	Exposure number per 1 ug/m3 category
	Road traffic	Laeq, Lden or Lnight in dB	Country/County/Municipality	Norway	2016	Exposure number per 5 dB band
	Railway	Laeq, Lden or Lnight in dB	Country/County/Municipality	Norway	2016	Exposure number per 5 dB band
	Aircraft	Laeq, Lden or Lnight in dB	Country/County/Municipality	Norway	2016 - 2021	Exposure number per 5 dB band
Portugal	PM _{2.5}	Annual mean (daily)concentration in µg/m ³	Municipality	Portugal	2021	Air quality monitoring stations
	O ₃	Annual mean of daily maximum 8-hour mean concentration in µg/m ³	Municipality	Portugal	2021	Air quality monitoring stations
	NO ₂	Annual mean concentration in µg/m ³	Municipality	Portugal	2021	Air quality monitoring stations
	Road traffic	Lden or Lnight in dB	Lisbon Metrop. Area, North	6 Portuguese cities	2016	Exposure number per 5 dB band
	Railway	Lden or Lnight in dB	Lisbon Metrop. Area, North	6 Portuguese cities	2016	Exposure number per 5 dB band
	Aircraft	Lden or Lnight in dB	Lisbon Metrop. Area, North	Lisbon and Porto	2016	Exposure number per 5 dB band
EU	PM _{2.5}	Annual mean concentration in µg/m ³	1km*1km	EEA38	2021	Gridded model output (regression-interpolation-merging mapping)
	O ₃	SOMO35 in ppb	1km*1km	EEA38	2021	Gridded model output (regression-interpolation-merging mapping)
	NO ₂	Annual mean concentration in µg/m ³	1km*1km	EEA38	2021	Gridded model output (regression-interpolation-merging mapping)
	Road traffic	Lden or Lnight in dB	Agglomeration/Region/Country	EEA38 (except Albania, Kosovo, Serbia)	2016	Exposure number per 5 dB band
	Railway	Lden or Lnight in dB	Agglomeration/Region/Country	EEA38 (except Albania, Kosovo, Serbia)	2016	Exposure number per 5 dB band
	Aircraft	Lden or Lnight in dB	Agglomeration/Region/Country	EEA38 (except Albania, Kosovo, Serbia)	2016	Exposure number per 5 dB band

4.3 Health outcomes

The health outcomes relevant to the study were retrieved from the work developed in WP1 regarding the possible risk-outcome pairs. The health outcomes included both the occurrence of the disease (prevalence and/or incidence) and number of deaths due to the diseases. The latter were defined according to the International Classification of Diseases - 10th revisions (ICD-10) (see **Table 4-3**). The codes were derived by the disease definitions used in the 2021 Global Burden of Disease study (GBD) [4]. Noise annoyance and sleep disturbance did not have ICD codes. For strokes, depending on the data available in the country, it was asked to give information on level 4 strokes (i.e. ischemic stroke, intracerebral hemorrhage and subarachnoid hemorrhage), otherwise to provide information on the level 3 stroke, following GBD definition. For each health outcome (i.e. incidence or prevalence of the disease and cause-specific mortality), it was asked to report:

- The geographical area for which the measurement is available (e.g. municipality, region, country level). It was possible to indicate up to three sources, the smallest areas or highest resolutions being the most important for the study;
- The reference period (e.g., year), or if the data was available for multiple periods, the most recent year;
- The definition used for the disease during the measurement of incidence, prevalence or mortality (example: ICD code, self-reported, medication consumption, other type of codes);
- The data input used (example: hospital discharge, national registries, survey, administrative data);
- The accessibility of the data, by choosing one of four options:
 - The data is owned by the institution of the user and can be used freely;
 - The data is owned by another institution but is available as open access;
 - The data is owned by another institution and can be used given authorisation of the owner (e.g., by signing a contract);
 - The data is owned by another institution and cannot be used.
- The source of the data (institution in charge) and a weblink to the data in case of open access;
- Whether health outcome numbers are stratified by age and/or sex.

See **Appendix 2.3** (prevalence) and **appendix 2.4** (mortality) for an example.



Table 4-3: Definition of diseases using ICD codes

Disease level 2 (according to GBD)	Disease level 3 (according to GBD)	Disease level 4 (according to GBD)	ICD code
Neoplasm	Lung cancer		ICD10: C33, C34–C34.92, Z12.2, Z80.1–Z80.2, Z85.1–Z85.20
	Malignant neoplasm of bladder		ICD10: C67
	Malignant neoplasm of kidney, except renal pelvis		ICD10: C64
	Skin cancer (non-melanoma)		ICD10: C44
Cardiovascular diseases	Ischemic heart disease		ICD10: I20–I25
		Myocardial infarction	ICD10: I21–I23
		Angina Pectoris	ICD10: I20
	Stroke		ICD10: G45–G46.8, I60–I63.9, I65–I66.9, I67.0–I67.3, I67.5–I67.6, I68.1–I68.2, I69.0–I69.3
		Ischemic stroke	G45–G46.8, I63–I63.9, I65–I66.9, I67.2–I67.3, I67.5–I67.6, I69.3
		Intracerebral hemorrhage	I61–I62, I62.1–I62.9, I68.1–I68.2, I69.1–I69.2
		Subarachnoid hemorrhage	I60–I60.9, I62.0, I67.0–I67.1, I69.0
		Heart failure, unspecified	ICD10: I50. 9
	Essential (primary) hypertension		ICD10: I10
Diabetes, urogenital, blood, and endocrine diseases	Type II Diabetes		ICD-10: E11.2, E11.21, E11.22, E11.29
	Chronic kidney disease		ICD10: N18
Chronic respiratory diseases	Chronic Obstructive Pulmonary Disease		ICD 10: J41, J42, J43, J44, and J47
	Asthma		ICD10: J45 and J46
Mental disorders	Depressive episode		ICD10: F32
Neonatal disorder	Disorders of newborn related to slow fetal growth and fetal malnutrition	Newborn small for gestational age	ICD10: P05.1
	Overweight, obesity and other hyperalimentation	Obesity	ICD10: E66.0
	Disorders of bone density and structure	Osteoporosis	ICD10: M80, M81, M82
	Disturbance of activity and attention		ICD10: F90.0



4.3.1.1 EU availability

a) Disease prevalence and/or incidence

At European scale, data on the prevalence and incidence (see text box 3) of selected diseases are available at NUTS2 level for the 2021 reference period (**Table 4-4**). Several, however, are unavailable, as is the case for sleep disturbance, annoyance, cognitive impairment, overweight and obesity, disorders of bone density and structure, disorders of newborns, disturbance and attention and kidney malignant neoplasms. Data represent in-patients (i.e. treatment that requires a patient to be admitted to a hospital or other care facility for at least one night) and out-patients (i.e. a patient or case that comes into hospital for treatment or care and is dealt with and released in the course of one day), collected according to hospital discharge data.

All data concerning the prevalence, incidence and mortality of the diseases selected at European level come from [Eurostat databases](#).

Text box 3 – Prevalence is the proportion of people affected by a disease at a specific time point, and **incidence** is the rate of new cases for a specific time range.

b) Disease mortality

As with disease prevalence and incidence, mortality rates for the diseases selected are available at NUTS2 level for Europe for the 2021 reference period (Table 4-5). Mortality on specific types of strokes, heart failure, disorders of newborns and skin cancer malignant neoplasms were not publicly available through the Eurostat databases. Data were collected based on national mortality registries.

4.3.1.2 Case country availability

a) Disease prevalence and/or incidence

The prevalence and/or incidence of selected diseases are available in most of the five case countries, and on a high geographical resolution (department, municipality, county or parish) (Table 4-4). However, some data are unavailable in certain countries. France, Norway and Portugal are the only countries to have data for specific types of strokes, cognitive impairment and disturbance of activity and attention.

Prevalence, incidence and mortality data for the selected diseases come from various sources depending on the country, such as (cancer) registries, hospital discharge data, hospitalization and patient admission data, health interview survey, health reimbursement data, medical bill, diagnose chart and mortality registries.

b) Disease mortality

Some non-fatal diseases are not taken into account here because they are not the underlying cause of death, such as sleep disturbance, annoyance, cognitive impairment, hypertension, overweight and obesity, disturbance of activity and attention, disorders of bones density and depressive episode. Mortality data for selected diseases is also available at a high geographical resolution for all case countries (from department to statistical sectors), with the



exception of mortality data for neoplasms, which are only available at country level in Belgium (**Table 4-5**).

Table 4-4: Data availability of disease prevalence and/or incidence. The disease level, smallest geographical area and last available reference period are displayed by country.

		EU		Belgium		Estonia	
Disease level 3	Disease level 4	Geographical area	Reference period	Geographical area	Reference period	Geographical area	Reference period
Lung cancer		NUTS 2	2021	Region	2021	Municipality	2022
Ischemic Heart Disease		NUTS 2	2021	Region	2020	Municipality	2022
Stroke		NUTS 2	2021	Province	2018	Municipality	2022
	Ischemic stroke	-	-	-	-	-	-
	Intracerebral hemorrhage	-	-	-	-	-	-
	Subarachnoid hemorrhage	-	-	-	-	-	-
Type II Diabetes		NUTS 2	2021	Province	2020	Municipality	2022
Chronic Obstructive Pulmonary Disease		NUTS 2	2021	Province	2018	Municipality	2022
Asthma		NUTS 2	2021	Province	2018	Municipality	2022
Heart failure, unspecified		NUTS 2	2021	-	-	Municipality	2022
Sleep disturbance		-	-	Province	2018	Municipality	2019
Annoyance		-	-	Province	2018	3 counties	2015
Cognitive impairment		-	-	-	-	-	-
Essential (primary) hypertension		NUTS 2	2021	Region	2018	Municipality	2022
Overweight, obesity and other hyperalimentation	Obesity	Country	2019	Region	2018	Municipality	2022
Disorders of bone density and structure	Osteoporosis	-	-	Region	2018	-	-
Disorders of newborn related to slow fetal growth and fetal malnutrition	Newborn small for gestational age	-	-	Region	2022	Municipality	2022
Chronic kidney disease		NUTS 2	2021	Region	2021	Municipality	2022
Disturbance of activity and attention		-	-	-	-	-	-
Depressive episode		NUTS 2	2021	Region	2018	Municipality	2022
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue	Bladder	NUTS 2	2021	Region	2022	Municipality	2022
	Kidney, except renal pelvis	-	-	Region	2022	Municipality	2022
	Skin cancer (non-melanoma)	NUTS 2	2021	Region	2022	Municipality	2022



Table 4-4 (continuation – France, Norway and Portugal)

		France		Norway		Portugal	
Disease level 3	Disease level 4	Geographical area	Reference period	Geographical area	Reference period	Geographical area	Reference period
Lung cancer		Department	2021	County	2022	Parish	2018
Ischemic Heart Disease		Department	2021	County	2021	Parish	2018
Stroke		Department	2021	County	2021	Parish	2018
	Ischemic stroke	County	2017	County	2021	Parish	2018
	Intracerebral hemorrhage	County	2017	County	2021	Parish	2018
	Subarachnoid hemorrhage	County	2017	County	2021	Parish	2018
Type II Diabetes		Department	2021	County	2021	Parish	2018
Chronic Obstructive Pulmonary Disease		Department	2021	County	2021	Parish	2018
Asthma		Department	2021	County	2021	Parish	2018
Heart failure, unspecified		Department	2019	County	2021	Parish	2018
Sleep disturbance		-	-	County	2022	Parish	2018
Annoyance		-	-	County	2021	Parish	2018
Cognitive impairment		Department	2021	County	2021	Parish	2018
Essential (primary) hypertension		Region	2021	County	2021	Parish	2018
Overweight, obesity and other hyperalimentation	Obesity	Department	2020	County	2021	Parish	2018
Disorders of bone density and structure	Osteoporosis	-	-	County	2021	Parish	2018
Disorders of newborns related to slow fetal growth and fetal malnutrition		Department	2013	County	2021	Parish	2018
Chronic kidney disease		Department	2021	County	2021	Parish	2018
Disturbance of activity and attention		Department	2021	County	2021	Parish	2018
Depressive episode		Department	2021	County	2021	Parish	2018
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue	Malignant neoplasm of bladder	Department	2021	County	2022	Parish	2018
	Malignant neoplasm of kidney, except renal pelvis	Department	2021	County	2022	Parish	2018
	Skin cancer (non-melanoma)	Department	2021	County	2022	Parish	2018

**Table 4-5 : Data availability of disease mortality (EU, Belgium and Estonia).** *The disease level, smallest geographical area, and last available reference period are displayed.*

		EU		Belgium		Estonia	
Disease level 3	Disease level 4	Geographical area	Reference period	Geographical area	Reference period	Geographical area	Reference period
Lung cancer		NUTS2	2020	Statistical sector	2020	Municipality	2022
Ischemic Heart Disease		NUTS2	2020	Statistical sector	2020	Municipality	2022
Stroke		NUTS2	2020	Region	2020	Municipality	2022
	Ischemic stroke	-	-	Statistical sector	2020	Municipality	2022
	Intracerebral hemorrhage	-	-	Statistical sector	2020	Municipality	2022
	Subarachnoid hemorrhage	-	-	Statistical sector	2020	Municipality	2022
Type II Diabetes		NUTS2	2020	Statistical sector	2020	Municipality	2022
Chronic Obstructive Pulmonary Disease		NUTS2	2020	Statistical sector	2020	Municipality	2022
Asthma		NUTS2	2020	Statistical sector	2020	Municipality	2022
Heart failure, unspecified		-	-	Statistical sector	2020	Municipality	2022
Disorders of newborn related to slow fetal growth and fetal malnutrition		-	-	-	-	Municipality	2022
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue	Bladder	NUTS2	2020	Country	2021	Municipality	2022
	Kidney, except renal pelvis	NUTS2	2020	Country	2021	Municipality	2022
	Skin cancer (non-melanoma)	-	-	Country	2021	Municipality	2022



Table 4-5 (continuation) : France, Norway and Portugal

		France		Norway		Portugal	
Disease level 3	Disease level 4	Geographical area	Reference period	Geographical area	Reference period	Geographical area	Reference period
Lung cancer		Department	2020	County	2022	Municipality	2021
Ischemic Heart Disease		Department	2020	County	2022	Municipality	2021
Stroke		Department	2020	County	2022	Municipality	2021
	Ischemic stroke	-	2020	County	2022	Municipality	2021
	Intracerebral hemorrhage	-	2020	County	2022	Municipality	2021
	Subarachnoid hemorrhage	-	2020	County	2022	Municipality	2021
Type II Diabetes		Department	2020	County	2022	Municipality	2021
Chronic Obstructive Pulmonary Disease		Department	2020	County	2022	Municipality	2021
Asthma		Department	2020	County	2022	Municipality	2021
Heart failure, unspecified		Department	2020	County	2022	Municipality	2021
Disorders of newborn related to slow fetal growth and fetal malnutrition		Department	2020	County	2022	Municipality	2021
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue	Malignant neoplasm of bladder	Department	2020	County	2022	Municipality	2021
	Malignant neoplasm of kidney, except renal pelvis	Department	2020	County	2022	-	2021
	Skin cancer (non-melanoma)	Department	2020	County	2022	-	2021

4.4 Deprivation

Sub-national, national, and European data sources were screened for data regarding the indicators selected within task 3.1 of the BEST-COST project. These indicators will be combined in a multiple deprivation index taking into account cultural and temporal validity for the case countries and the EU countries in general. Given that the focus of BEST-COST and, specifically WP3, was to prioritise a high geographical resolution, the final list of indicators was also assessed based on the availability of these at small geographical scales, defined for these purposes at the municipality level (i.e., LAU 2) at least. The following five indicators fulfilled these criteria :

- percentage of households without central heating,
- percentage of high school graduates in the population,
- percentage of unemployed individuals in the active population (aged 18-65 years),
- percentage of a single-parent households, and
- percentage change of population size over the previous 5 years.

These indicators would cover five domains (basic amenities, education, employment, family structure, and demographics).

For data collection, partners in the five case countries were asked to provide the following information for each indicator:

- The description of the indicator (example for “high school graduates in the population”: definition for Belgium = “percentage of high school graduates in the population”);
- The geographical area for which the measurement is available (e.g. municipality, region, country level). It was possible to indicate up to three sources, the smallest areas or highest resolutions being the most important for the study;
- The latest year for which the data is available;
- The type of the data (example: census, survey, registry,...);
- The accessibility of the data, by choosing one of four options:
 - The data is owned by the institution of the user and can be used freely;
 - The data is owned by another institution but is available as open access;
 - The data is owned by another institution and can be used given authorization; of the owner (e.g., by signing a contract);
 - The data is owned by another institution, and cannot be used.
- Web link where data can be found;
- Whether the indicator is stratified by age and/or sex.



4.4.1 EU and case country availability

Table 4-6 shows the smallest geographical level for which the indicators are (un)available in the case study countries and all European countries.

At a broader geographical scale (i.e., NUTS level-2 regions), several of the above indicators are available via Eurostat (*Europe* column in Table 4-6). This dataset is publicly available, regularly updated, and harmonised, and available for most EU and the EEA countries. Eurostat also develops several socioeconomic indicators that could be combined to fulfil the requirements of the current project, such as persons who are at risk of poverty or severely materially deprived or living in households with very low work intensity [14]. However, it should be noted that not all data are available for each country and for each year.

We can see that the definition of the indicator may vary slightly from one country to another. For each country, data is available at a high geographical resolution (for Belgium, from the statistical sector to NUTS2; for Estonia, from the county to the municipality; for France, at most the municipal level; and for Portugal, from LAU2 to NUT2). For Norway, four indicators have data available at municipality level. Data for the “heating” indicator is missing at the moment.

Table 4-6: Data availability for the five selected deprivation indicators in the five countries concerned (Belgium, Estonia, France, Norway, Portugal) and in the European Union. For each country, the indicator is described and the lowest geographical level at which it is available is indicated.

		Belgium		Estonia		France	
Domain	Indicator	Description of the indicator	Geographical area	Description of the indicator	Geographical area	Description of the indicator	Geographical area
Basic amenities	Heating	Percentage of households without central heating	Province (NUTS2)	Percentage of households without central heating	County (LAU 1)	Number of main residences with collective central heating + number of main residences with individual central heating	IRIS
Education	School graduates	Percentage of high school graduates in the population	Statistical sector	Percentage of high school graduates in the population	Municipality (LAU2)	Number of unschooled people aged 15 or over with a Baccalaureate, professional certificate or equivalent	IRIS
Employment	Unemployment individuals	Percentage of unemployed individuals in the active population (aged 18-65)	Statistical sector	Percentage of unemployed individuals in the active population (aged 16–60)	Municipality (LAU 2)	Number of unemployed aged 15 to 64	IRIS
Family structure	Single-parent household	Percentage of a single-parent households	Statistical sector	Percentage of a single-parent households	Municipality (LAU 2)	Number of single-parent families	IRIS
Demographics	Change in population	Percentage change of population size over the previous 5 years	Municipality	Percentage change of population over the previous 5 years	Municipality (LAU 2)	Percentage of population change over the period 2014-2020	Municipality



Table 4- 6: continuation

		Norway		Portugal		Europe	
Domain	Indicator	Description of the indicator	Geographical area	Description of the indicator	Geographical area	Description of the indicator	Geographical area
Basic amenities	Heating	Not available		Private households in conventional dwellings of usual residence by place of residence and type of heating most often used	Parish (LAU 2)	Percentage of households without central heating	NUTS 2
Education	School graduates	Percentage of high school graduates in the population	Municipality	Percentage of high school graduates in the population	Parish (LAU 2)	Percentage of high school graduates in the population	NUTS 2
Employment	Unemployment individuals	Percentage of unemployed individuals in the active population (aged 16-60)	Municipality	Unemployment rate in the population aged 16-89	Parish (LAU 2)	Percentage of unemployed individuals in the active population (aged 16-60)	NUTS 2
Family structure	Single-parent household	Single mother/father with young children	Municipality	Proportion of mono-parental family nuclei	Parish (LAU 2)	Percentage of a single-parent households	NUTS 3
Demographics	Change in population	Percentage change of population size over the previous 5 years	Municipality	Resident population by Place of residence	Administrative regions (NUTS 2)	Percentage change of population over the previous 5 years	NUTS 2

4.5 Summary and highest common resolution

This chapter aimed to provide some insights into the data availability in the five case study countries and at EU level regarding environmental exposures, health outcomes and socioeconomic deprivation. A particular focus was put on the highest geographical resolution at which data is available for all case study countries, which was the focus of the work in task 3.2.

In general, we identified great variability in data availability in terms of geographical resolution from country to country (see Table 4-7). Nevertheless, for most indicators we were able to present a common resolution for data analysis. For air pollution, the best agreement was found in the EU data, which are available at a 1km² scale for the overall EU territory. Finding a common resolution was more complex for noise pollution, where data are accessible at lower geographical resolutions. For example, for EU this is only available at the country level. However, some countries such as Norway and Portugal have data available on a higher geographical resolution.

For disease prevalence/incidence, the resolution usually depends on the disease of interest. The data are available at regional level (NUTS 2) for the EU, but most of the countries can provide higher resolution data (France, Estonia, Norway, Portugal). For disease mortality, it is possible to obtain data on a NUTS 3 scale, as each country individually presents data with this geographical scale.

The resolution for deprivation usually also depends on the indicator of interest. The data are available at regional level (NUTS 2) for the EU, but some countries (France, Estonia, and Norway) can provide higher-resolution data. Considering the composite nature of the deprivation index, all indicators need to be available at the same geographical level within one country.

Table 4-7: Summary table for pollutant exposure and disease outcome considering the highest geographical resolution available for each country and the EU.

	Pollutant exposure		Outcome		
	Air pollutant	Noise pollutant	Disease prevalence/incidence	Disease mortality	Deprivation
Belgium	10m * 10m	NUTS 1	NUTS 1 / NUTS 2	Statistical sector	NUTS2
France	4km * 4km	-	NUTS 3	NUT3	LAU2
Estonia	-	NUTS 1	LAU 2	LAU 2	LAU1
Norway	6km * 6km	NUTS 5 / LAU 2	NUTS 3	NUTS 3	LAU2
Portugal	8km * 8km	NUTS 3	LAU 2	LAU 1	NUTS2
EU	1km * 1km	Country	NUTS 2	NUTS 2	NUTS2



5 Methodological framework for data mapping

5.1 Outline of the proposed framework

Mapping is the mainstay of Cartography, a field that occupies itself with the design and production of - typically 2D - symbolized graphical representations of geographic areas, phenomena, and processes. Contemporary maps are generally made digitally using **GIS (Text box 3)** software, examples of which include QGIS, ArcGIS and AutoCAD Map. There is a sound scientific basis and a set of derived principles to guide map making [15], [16]. Yet it should be noted that map design requires some methodological and artistic choices to be made considering the target audience, the message to be conveyed by the maps, the medium of distribution and personal preference. Consequently, it is not possible to identify singularly optimal mapping solutions for any stated problem, as there are always alternate but equivalent approaches.

Text box 3 - Geoinformation Systems (GIS) are computer-based tools that analyse, visualize, and interpret geographical data. They integrate various data types, such as maps, satellite imagery, and statistical data, to provide insights into spatial patterns and relationships. GIS applications span multiple fields, including urban planning, environmental management, transportation, and public health. By enabling the visualization of spatial information and the performance of complex spatial analyses, GIS aids in decision-making, resource management, and problem-solving.

The methodological framework in this report does not cover all possible mapping techniques and designs. Instead, it focuses on proposing practical solutions for what we identified as the most common mapping needs that may present themselves to researchers working on the intersection of environmental science, social sciences, and epidemiology. In short, these needs entail comprehensible self-contained thematic maps, possibly complemented with support material, that allow the reader to visually grasp the spatial-quantitative distributions and interrelations between pertinent indicators reflecting environmental stressors, health outcomes, socioeconomic deprivation, ideally at the highest geographical resolution possible. Considering the variable resolutions at which the statistics of these topics are distributed within the EU countries and institutions, a challenge that was highlighted extensively in the previous chapter of this report, we decided to define five data availability scenarios to structure our framework. These scenarios are conceived along the lines of data dimensionality, i.e., the number of indicators under consideration within a single map or map series, and data resolution, i.e., the size of the geographical units on which the indicator statistics are made available. The five selected mapping scenarios are:

1. Mapping a single high-resolution indicator.
2. Mapping a single low resolution indicator.
3. Mapping two equal resolution indicators.
4. Mapping a higher and lower resolution indicator.
5. Mapping three equal resolution indicators.

While there is no convention as to what constitutes a high or low spatial resolution, it was decided, for reference and considering the typical resolutions of the datasets treated in this report, to label NUTS1- as lower resolution and NUTS2+ as higher resolution. In any case, this



definition is only indicative. Spatial data distributed in raster format with pixel sizes of 1km or smaller are also considered having a higher resolution for the purpose of this report.

The next section of this chapter summarizes the overarching mapping protocols that are applied throughout the development and illustration of the methodological framework. The following sections handle the proposed mapping techniques for each of the five data availability scenarios. Depending on the scenario there may be several complementary techniques to be discussed. The data used to illustrate the proposed mapping techniques mainly serve an exemplary purpose, but they also provide an opportunity to illustrate some of the findings of the previous chapter. The example maps also touch upon the type of analyses that may be of interest to the wider BEST-COST consortium.

5.2 General mapping protocol

5.2.1 Software and data

All maps produced for this report are made with open-source software. The workflows and scripts were implemented using QGIS, Python and SQL. Other scripting languages, like R, are equally suitable for this task. As said in the introduction, the developed scripts are shared through the BEST-COST GitHub portal, and the technical description will be handled via separate BEST-COST protocols.

Likewise, for transparency and reproducibility, most example data used to illustrate the mapping scenarios are obtained from publicly available sources. If protected data were used to produce maps, this was clearly indicated. The used example data is clarified at the start of the concerned data availability scenario.

The spatial definitions of the mapped geographical units, that include NUTS regions [12] and world countries, with separate definitions of land and sea borders [17], were obtained from EUROSTAT. This data is used in nearly every map produced for this report. Unless stated otherwise, we use the most recent version of the NUTS region definitions, which currently date from 2021.

5.2.2 Support graphs

Maps excel at illustrating the distribution of data within a geographical space, yet they do not always succeed in conveying the statistical qualities of the dataset under consideration. That is why we decided to also provide one or more support graphs for each data availability scenario. The support graphs provide an alternate and complementary visualization of the mapped indicators' (co-)distribution in abstract data space, rather than geographical space.

5.2.3 Data processing

Additional data processing on top of the processing already performed by the data providers is avoided to simplify the mapping workflows. Seeing however that published statistical datasets are often affected by missing data, some degree of **data imputation (Text box 4)** may have to be applied to produce more complete maps. For simplicity, we first try temporal or historic imputation, i.e., replacing a missing value with the most recent available value prior to the period of the missing value. An alternate form of imputation for spatial data, that may be



applied if required, is aggregation. In this approach, missing data for one or several smaller regions within a larger region are addressed by only mapping the available data on the larger region. If data imputation is used, this will be indicated in the map and the corresponding discussion.

Text box 4 - Data imputation is the process of replacing missing or incomplete data with substituted values to maintain dataset integrity and enable accurate analysis. Methods include statistical techniques like mean, median, or mode substitution, regression imputation, and more sophisticated approaches like multiple imputation or machine learning algorithms. Imputation helps in retaining the full dataset size, improving the performance of analytical models, and ensuring robust and reliable insights. Proper imputation is crucial for reducing bias and preserving the original data's structure and relationships.

Where needed, the mapped indicators are surface-, population-, and/or age-normalized to facilitate comparison of values between regions or countries with varying geometries and demographics. Again, this will be indicated in the maps and discussions where appropriate. To further facilitate map interpretation, continuous quantitative data are typically discretised into a small number of ordered classes, that each entail a bounded value interval. Several methods can be used to achieve this classification, depending on the underlying data distributions and/or mapping requirements [16]. Within our work, it was decided to use either domain-specific manual classification or the widely used **quantiles classification** [18] (**Text box 5**) method.

Text box 5 - Quantiles classification is a method of dividing data into equal-sized intervals based on percentiles, such as quartiles, quintiles, or deciles. Each class contains an equal number of data points, making it useful for visualizing and comparing distributions across different datasets. This method is particularly beneficial to help in identifying patterns and disparities by ensuring each class is equally represented. It simplifies interpretation and highlights relative standings within the data, facilitating informed decision-making and communication of complex information.

5.2.4 Map design guidelines

The map design used for the proposed methodological framework follows the best current practices, to the extent allowed by the parameters of this assignment [15], [19]. A thorough exploration of these practices also falls outside the scope of this report. Suffice to say that the principles of visual balance as well as **intellectual and visual hierarchy** (**Text box 6**) are to be applied, while making the maps as self-explanatory and self-containing as possible. For the sake of inclusivity, efforts were also made to use colourblind-friendly colour schemes for the produced maps and graphs, or to provide additional colourblind-friendly alternatives. For this purpose, we draw on ColorBrewer¹, an excellent source of perceptually optimized and colourblind-friendly colour schemes. ColorBrewer schemes are by default included in QGIS and Python modules like Matplotlib.

¹ <https://colorbrewer2.org/>



Text box 6 - Intellectual and visual hierarchy organize map elements by importance to enhance clarity and usability. Intellectual hierarchy involves prioritizing information based on relevance and purpose, guiding the map's design and content selection. Visual hierarchy applies design principles like size, colour, contrast, and placement to emphasize key features and ensure readability. Important elements, such as main geographic features or critical data, are made more prominent, while less critical information is subdued. Together, these hierarchies help users to quickly interpret and navigate maps.

5.2.5 Map extent and Coordinate Reference Systems

The maps made for this report cover the EU and its Candidate Countries (CC). We use the Coordinate Reference System (CRS) called *European Terrestrial Reference System 1989 Lambert Azimuthal Equal Area* (ETRS89-LAEA)², that is widely used for maps covering on- and offshore Europe. Inset maps showing smaller regions or overseas European territories may deviate from this CRS. More information on coordinate systems and **map projections** (**Text box 7**) can be obtained from Slocum, T.A et al. [16].

Text box 7 - Map projections are methods of representing the curved surface of the Earth on a flat plane. Because the Earth is a sphere, no projection can preserve all geographic properties (shape, area, distance, and direction) simultaneously, leading to inevitable distortions. Different projections prioritize different properties: for instance, the Mercator projection preserves direction, making it useful for navigation, while the Albers equal-area projection maintains area proportions, useful for statistical maps. The choice of projection depends on the map's purpose, as each has trade-offs.

5.3 Mapping approach per data availability scenario

5.3.1 Scenario 1 – Mapping a single high-resolution indicator

5.3.1.1 Selected example data

The example topic selected for this data availability scenario is NO₂ air pollution within the EU and its Candidate Countries (CC). As with O₃ and PM_{2.5}, interpolated NO₂ air pollution data are made available in raster format by the EEA, quantified in µg/m³ at yearly intervals and on a relatively high spatial resolution of 1km². For this map, we took the currently most recent available data of 2021 [20]. To assess the population exposed to NO₂ air pollution, we draw on the EUROSTAT Census grid 2021 (version 13 March 2023) that is also published on a 1km² resolution [21].

5.3.1.2 Mapping

Because we're mapping quantitative thematic data at a high spatial resolution, especially relative to the geographic extent of the mapped area, we judge that **choropleth mapping** (**Text box 8**) is the most suitable approach here. The map illustrating this scenario is shown in Figure . We've decided to use domain-specific knowledge to discretise the continuous pollution data into five classes. The highest pollution class covers average yearly NO₂ air concentrations exceeding the WHO threshold of 10 µg/m³ [22]. The lower pollution classes were defined at decreasing 2.5 µg/m³ [22] intervals, for ease of interpretation. While the definition of what constitutes problematic pollution is up for discussion, this threshold is mainly used here to

² <https://epsg.io/3035>



provide some context for numbers that may otherwise be hard to interpret by laymen. Other threshold values or classification systems derived from legislation, e.g., the EU Air Quality Directive, [23] could also have been used here.

Text box 8 - Choropleth mapping is a cartographic technique to represent statistical data through varying shades or colours within predefined areas, such as countries, states, or districts. Each colour or shade corresponds to a data range, effectively visualizing spatial distributions and patterns. This method allows for quick and intuitive comparison of data across regions, highlights spatial trends, and is easy to interpret for a wide audience. Choropleth maps make complex datasets accessible and aid in data-driven decision-making.

Interpolated NO₂ air pollution

EU27 + Candidate countries + UK - 2021 - 1km² grid - Manual classification based on WHO guidelines

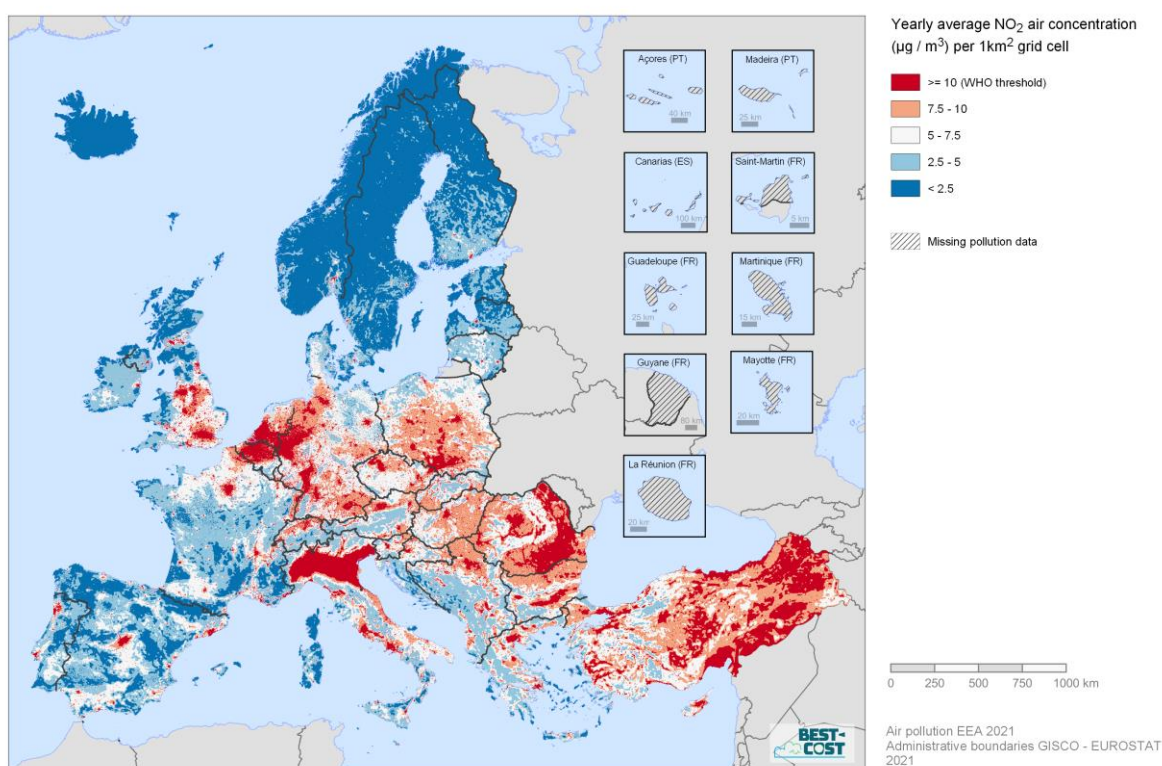


Figure 5-1. Choropleth map with a bivariate colour scheme showing 2021 NO₂ air pollution on 1km² resolution in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#).

NO₂ air concentrations above the WHO threshold are considered problematic and may entail substantial public health risks for the population exposed. To enhance this message in the map, we've decided to use a blue-red **bipolar colour scheme** (Text box 9), with red shades for the higher and more problematic air pollution classes and blue shades for the lower and less problematic pollution classes. A **unipolar colour scheme** (Text box 9), for instance only showing shades of red, might also have been used, but such a scheme would create less visual contrast between the lower and higher pollution classes. A red-blue bipolar colour scheme is considered a colourblind-friendly option.



Text box 9 - Unipolar colour schemes use a single colour gradient, ranging from light to dark or from one shade to another, to represent increasing or decreasing values of a single variable. They are ideal for visualizing data with a natural progression, such as population density or temperature. **Bipolar colour schemes** employ two contrasting colours that meet at a neutral midpoint, representing a divergence from a central value. They are effective for displaying data with positive and negative values, such as temperature anomalies or changes from a baseline. The contrast between the two colours highlights deviations, making it easier to interpret data trends and patterns.

Note that country borders were plotted over the pollution raster to provide some geographic context to the mapped data. Overseas European territories, including among others the Canary Islands, Madeira, and Saint-Martin, were included in the map through smaller **inset maps (Text box 10)** arranged on the right of the main map. As is common when mapping spatial statistics, there are some regions that do not have data. In this map, we decided to mark these regions with a grey hatched pattern fill, to distinguish them from regions with data without drawing too much visual attention.

Text box 10 - Inset maps are smaller maps set within the larger main map. They serve several purposes: providing a zoomed-in view of a specific area, showing locations too distant to fit the main map, highlighting details of complex regions, or displaying thematic information at a different scale. Insets enhance map readability and functionality by offering additional context or focusing on areas of interest without cluttering the main map.

Air pollution is a topic of interest for BEST-COST and it is relevant to have an idea of how populations are exposed to air pollution seeing that this would be determining for health outcomes like morbidity and mortality. To clarify how the mapping proposed in this scenario can also support the visualisation of pollution exposure, we produced the map shown in Figure . This map displays on NUTS3-level the percentage of the population exposed to a yearly average NO₂ air pollution exceeding the WHO threshold. This percentage was obtained by first multiplying the population grid data with a binary mask identifying cells with NO₂ air pollution above threshold and then normalizing the resulting NUT3-level sums with the corresponding total population. The mapped classes were obtained here using quantiles classification (**Text box 5**) with rounding. A bipolar colour scheme (**Text box 9**) was again used in this map, but with opposing purple and orange hues, with the latter indicating higher exposure and vice versa.

**Population exposed to NO₂ air pollution**

EU27 + Candidate countries + UK - 2021 - NUTS3 - Quantiles classification

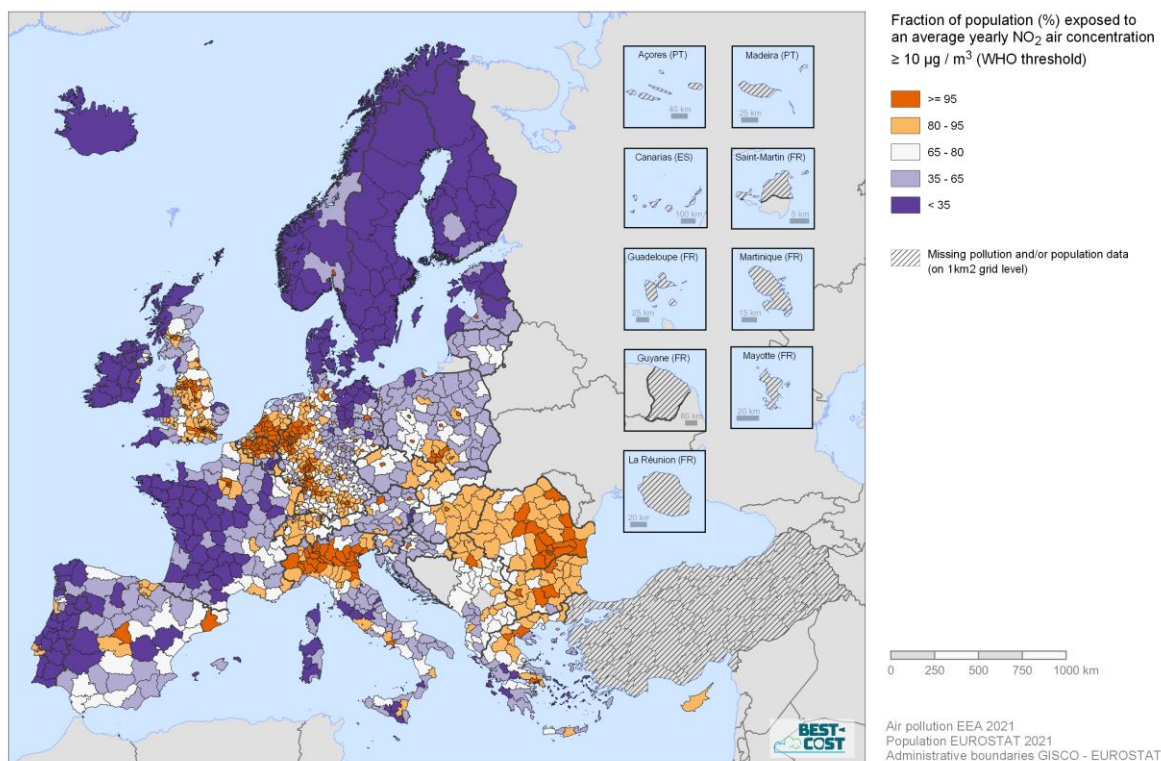


Figure 5-2. Choropleth map with a bivariate colour scheme showing 2021 NO₂ air pollution exposure on NUTS3-level in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#).

5.3.1.3 Support graph

Two types of support graphs are presented here to help interpret the data mapped in **Figure 5-2**. In **Figure 5-3** we show a colour-coded histogram displaying the underlying data distribution of the mapped NO₂ exposure. The height of each bar in this graph expresses the number of NUTS3 region with a certain percentage of the population, as denoted on the horizontal axis, living in an area with an average yearly NO₂ air concentration exceeding the WHO threshold. This representation may be more meaningful to assess health outcomes compared to reporting air pollution exposure in relative or absolute surface measures.

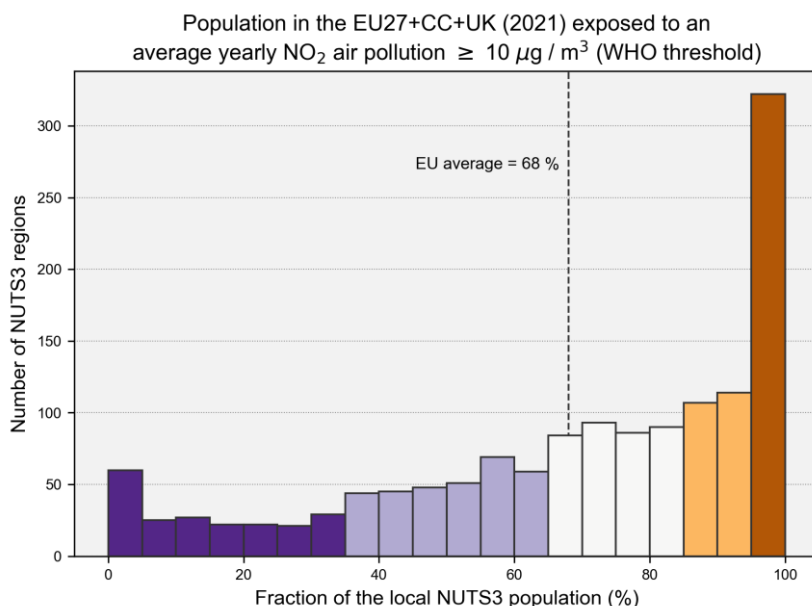


Figure 3-3. Colour coded histogram showing the quantitative distribution of the 2021 NO₂ air pollution exposure on NUTS3-level mapped in Figure 5-2. A high-resolution version of this figure can be consulted [here](#). CC = Candidate Countries.

Note that the colour coding used in the histogram matches the prior class and colour definitions used in Figure 5-2 to visually underline the linkage between the map and corresponding support graph. We also added a vertical dashed line in the support graph to show the EU average NO₂ exposure value, which provides an additional handle for the reader to process this information.

Secondly, we've produced a bar chart showing national NO₂ pollution exposure (Figure 5-4). The bars shown in this support graph are organized horizontally, meaning that the width of each bar represents the fraction of the corresponding national population living in an area with an above WHO threshold NO₂ air pollution [22]. The top-down organization highlights that this is a ranked representation. To clarify the exact positions of each country within the ranked data, we've added the matching indices between brackets after each country label on the vertical axis. At the top of the graph and separated from the rest of the bars, we show the EU average for reference. The bars showing the national data are again colour coded as in Figure 5-2, to clarify the linkage with the map.

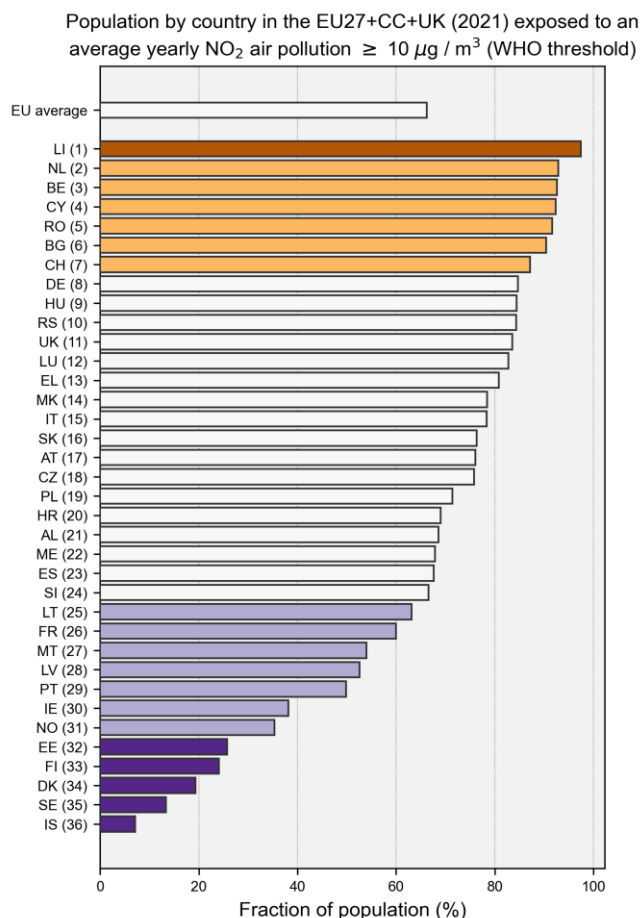


Figure 5-4. Colour coded horizontal bar chart showing 2021 national-level (NUTS0) NO₂ air pollution exposure in the EU27, its Candidate Countries and the UK. The bar colour coding matches that map layout of **Figure 5-2**. A high-resolution version of this figure can be consulted [here](#).

5.3.2 Scenario 2 – Mapping a single low resolution indicator

5.3.2.1 Selected example data

For this scenario we will look at the example of NUTS1-level mortality by disease/condition. To address the fact that mortality is strongly impacted by regional demographics, which may render interregional comparison unfeasible if crude death rates are used, we instead draw on age-standardized death rates, expressed as yearly deaths per 100 000 inhabitants. These standardized death rates are obtained by taking the weighted sum of local age cohort-specific death rates, using as weights the corresponding relative frequencies derived from the European Standard Population (ESP) edition 2013 [24]. Standardized mortality data are distributed by EUROSTAT under the data code *hlth_cd_asdr2* [25].

We use the most recent mortality statistics that are currently published, dating back to 2021. In the following maps we will focus on three disease/conditions groups: 1. malignant neoplasms (cancers), 2. diseases of the circulatory system and 3. other diseases. The former two condition groups are statistically identified in the EUROSTAT data using the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD10) [26], respectively using the condition codes *C0-97* and *I0-99*. The third condition group defined



above, i.e., other diseases, is defined as the union of all other condition groups without the former two.

5.3.2.2 Mapping

Considering the lower spatial resolution of the used geographic units, that display strong surface area and shape variations, we've decided that **proportional symbol mapping** (Text box 11) is a suitable approach for this data availability scenario. Note that several NUTS1 regions are identical to the matching country borders. First, we will look at the mapping of standardized mortality due to malignant neoplasm (Figure 5-5). Countries with reported mortality data are given a darker grey background, compared to the default lighter grey background, and on top of this darker background we plot circle diagrams with varying sizes and colours. The used colour scheme for the circle fill colours is of the unipolar type (Text box 9) with four hues ranging from white to dark red, the later indicating higher mortality and vice versa. Likewise, we've decided to also let the size of the circle diagrams reflect the magnitude of mortality, with larger sizes indicating higher mortality and vice versa. This dual distinction, while strictly speaking unnecessary, enhances the visual contrast between lower and higher classes and acts as an additional safeguard for colourblind users. The value ranges of the four classes were obtained using quantiles classification (Text box 5) with rounding for ease of interpretation. Note that we only use four circle diagram sizes for the classes, to facilitate visual distinction.

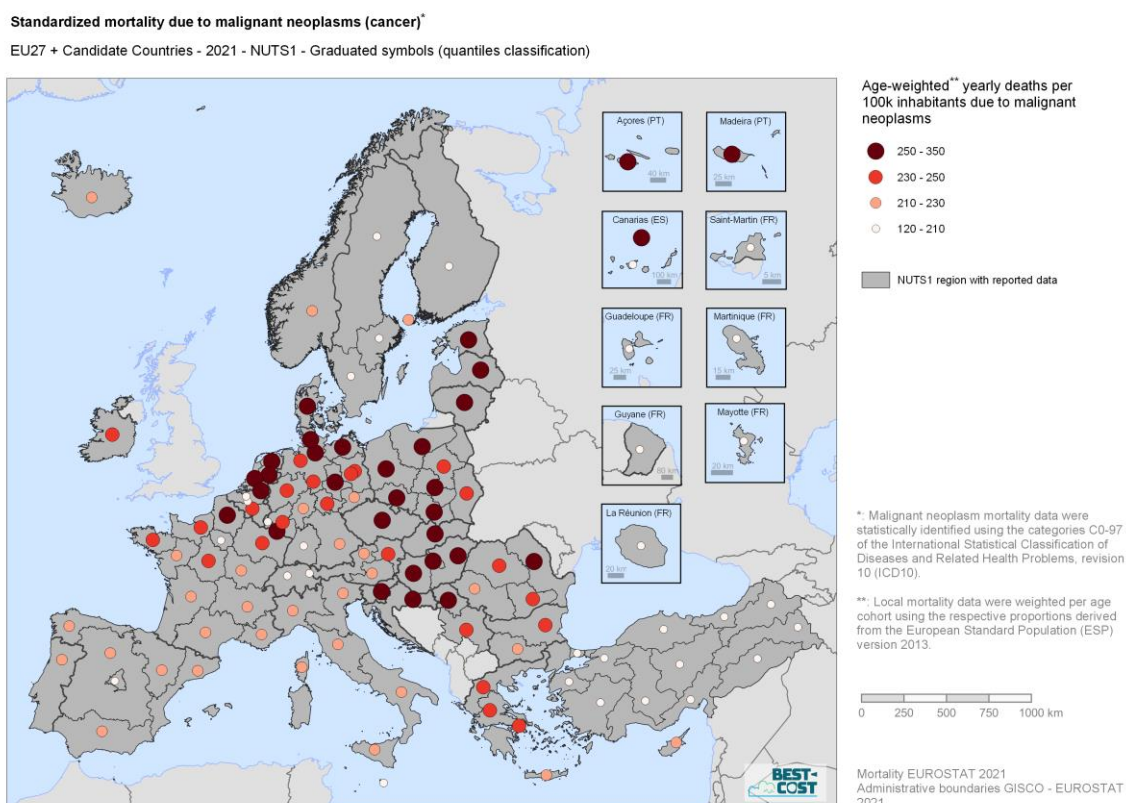


Figure 5-5. Proportionate symbol map, using circle diagrams, showing 2021 age-weighted mortality due to malignant neoplasms on NUTS1-level in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#).



Text box 11 - Proportional symbol mapping is a cartographic technique that uses symbols of varying sizes to represent data values, with larger symbols indicating higher values. This method is useful for visualizing precise quantities and comparing data across locations without being constrained by area size. However, it can become cluttered in maps with many smaller and densely concentrated geographic regions. Compared to choropleth mapping, proportional symbol mapping is less impacted by area bias, but it may lack the ease of interpretation and spatial coherence that choropleth maps provide.

Standardized mortality by cause of death*

EU27 + Candidate Countries - 2021 - NUTS1 - Scaled pie charts (quantiles classification)

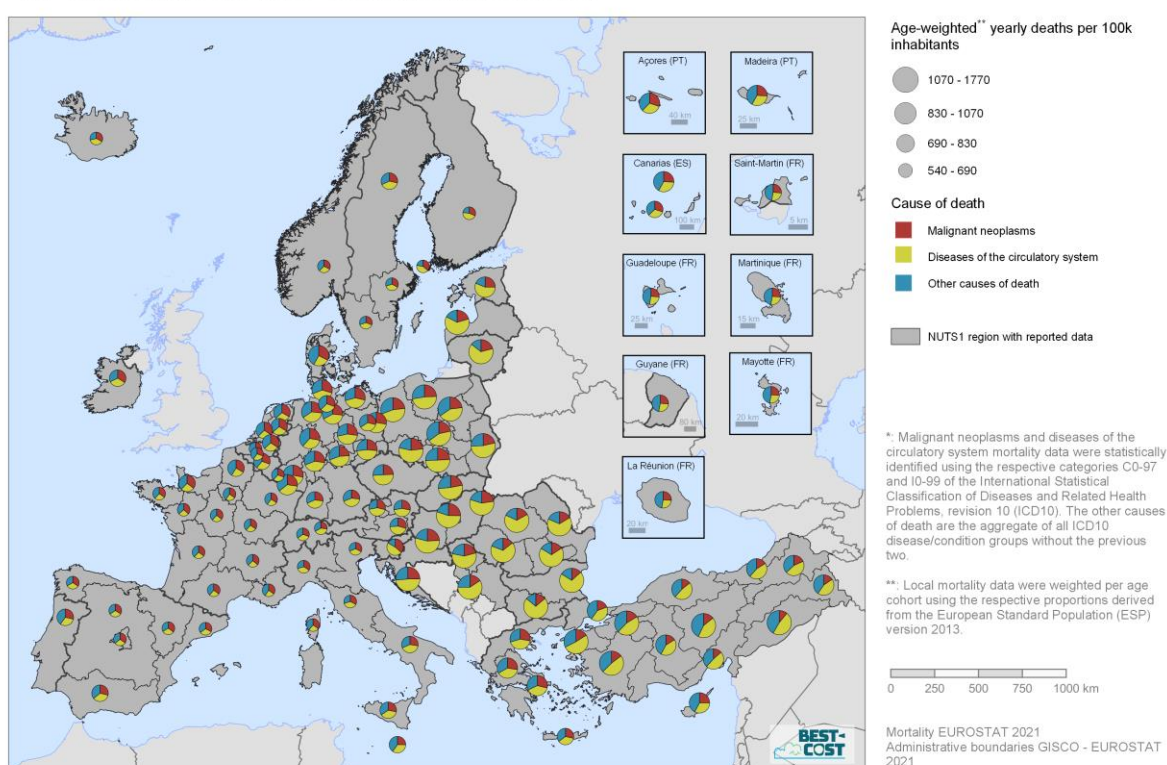


Figure 5-6. Pie chart map showing 2021 total age-weighted mortality and age-weighted mortality by cause of death on NUTS1-level in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#).

As an alternate and more advanced approach, we present a **pie chart mapping** (Text box 12) showing both the total standardized mortality and the composition of this total mortality with regards to the three disease/condition groups that were defined earlier (Figure 5-6). Total mortality is reflected by the circle diagram sizes, of which there are again four, and composition by the relative sizes of the colour-coded slices within each diagram. The colour coding used here is of a categorical nature because there is no logical order or rank associated with these classes. The value ranges of each pie size are again obtained with quantiles classification (Text box 5) and rounding.



Text box 12 - Pie chart mapping integrates pie charts into geographic maps, with each pie chart representing data for specific locations. Each slice of the pie shows the proportion of a particular variable, allowing multiple variables to be visualized simultaneously within a single geographic area. Advantages include the ability to convey multivariate data clearly and to compare parts to a whole directly on the map. However, disadvantages involve potential clutter and difficulty in interpreting overlapping or closely placed pie charts, especially in densely populated areas.

Note that some degree of diagram overlap is allowed in the pie chart map, as well as in the previous proportional symbol map, to cope with clusters of smaller regions. The overlap is particularly visible in Belgium and the Netherlands. Overlap is typically difficult to avoid all together in proportionate symbol and pie chart maps. A possible solution would be to displace overlapping diagrams and indicate the matching region with a leader line, but this approach may clutter the map and reduce its readability. Finding a solution for diagram overlap in any case requires a case-dependent assessment.

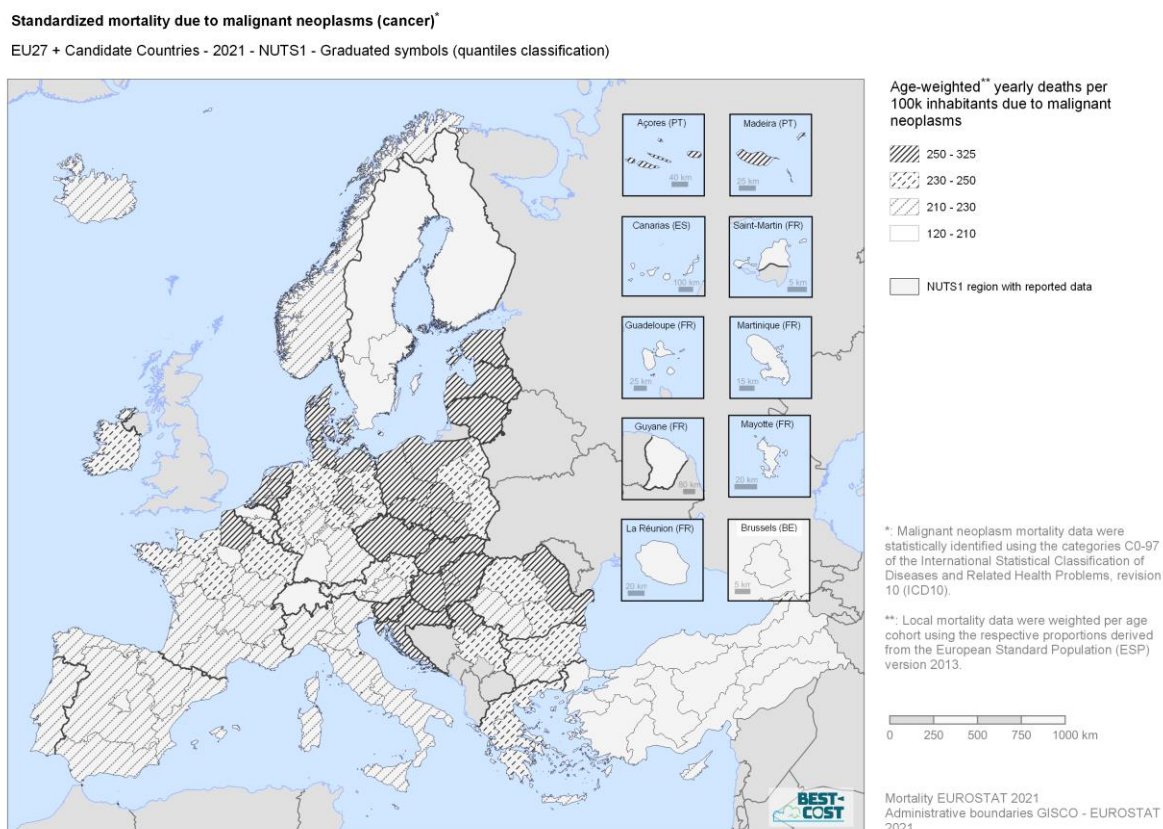


Figure 5-7. Line pattern map showing 2021 age-weighted mortality due to malignant neoplasms on NUTS1-level in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#).

In the third and final approach of this scenario, we provide an alternative for mapping a single indicator, for which we will retake the example of standardized mortality due to malignant neoplasms. Here we use **pattern fill mapping (Text box 13)** to distinguish different mortality magnitude classes (**Figure 5-7**). No fill indicates the lowest class, and the higher classes are marked by increasing densities of line pattern fills. Besides lowering the spacing between lines, we also



Text box 13 - Pattern fill mapping is a cartographic technique that uses various patterns, such as stripes, dots, or hatches, to fill areas on a map to represent different data values or categories. This technique can distinguish between multiple variables or highlight different qualitative data within the same geographic region. Advantages of pattern fill mapping include its ability to represent overlapping data and its effectiveness for colourblind users. Disadvantages include potential visual clutter and difficulty in distinguishing between similar patterns. Compared to colour-based methods like choropleth maps, pattern fill maps may be less intuitive and harder to read.

use drawing style, i.e., use of dashed or fully drawn line, as well as line thickness and colour to visually differentiate the densities of the mortality classes. Whereas line pattern mapping provides a notably colourblind-friendly alternative and can be combined relatively easily with other mapping techniques, it may by itself be less intuitive compared to proportionate symbols and choropleth maps.

Finally, we should point out that NUTS2 and lower resolution regions display large surface area variations. This makes it harder to discern the map layout of smaller regions in the map, especially if the map layout is contained within the region borders, as would be the case when using choropleth mapping and pattern fill mapping. In such cases, it may become necessary to include additional inset maps showing the map layout for smaller regions on an enhanced scale. To illustrate this point, we've added an additional inset map (**Text box 10**) for the region of Brussels in **Figure 5-7**, as well as in subsequent choropleth maps on NUTS2 or lower resolution.

5.3.2.3 Support graph

The support graphs proposed for this data availability scenario are elaborations of the earlier shown histogram and bar chart of the first scenario (**Figure 5-3 - Figure 5-4**). Continuing with the example of the standardized mortality data mapped in **Figure 5-6**, we plot a series of four smaller histograms to illustrate the quantitative distributions of total mortality, covering all causes of death, and mortality due to the more specific causes of death (**Figure 5-8**). The histogram bars of the specific causes of death follow the colour coding used in the map. For total mortality, which is reflected by pie chart size in the map, we use a neutral grey colour in the matching histogram. EU averages are plotted in their respective histograms to highlight the first moment of the value distributions.



Standardized mortality by cause of death in the EU27+CC in 2021

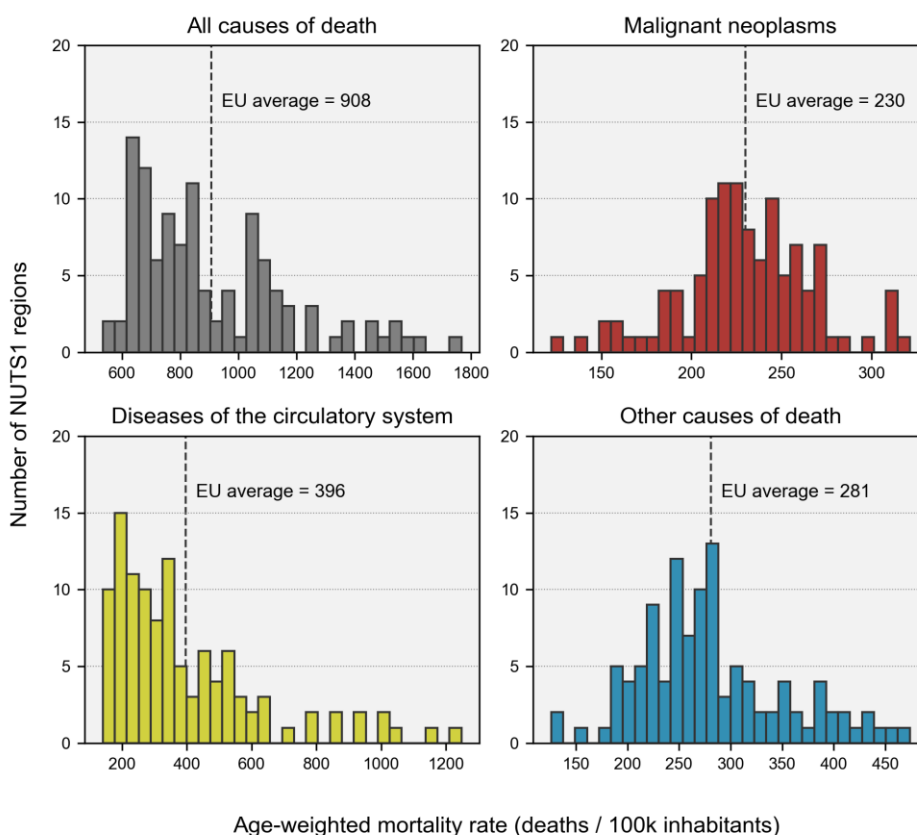


Figure 5-8. Histograms showing the quantitative distribution of 2021 age-weighted mortality due to all causes of death, malignant neoplasms, diseases of the circulatory system and other causes of death on NUTS1-level in the EU27 and its Candidate Countries. The colouring of the bars matches the respective class colours used in **Figure 5-6**. A high-resolution version of this figure can be consulted [here](#).

Also retaking the example of standardized mortality data by cause of death, the colour-coded bar chart shown in **Figure 5-9** illustrates the EU average and ranked national averages as well as the composition of these averages by cause of death.

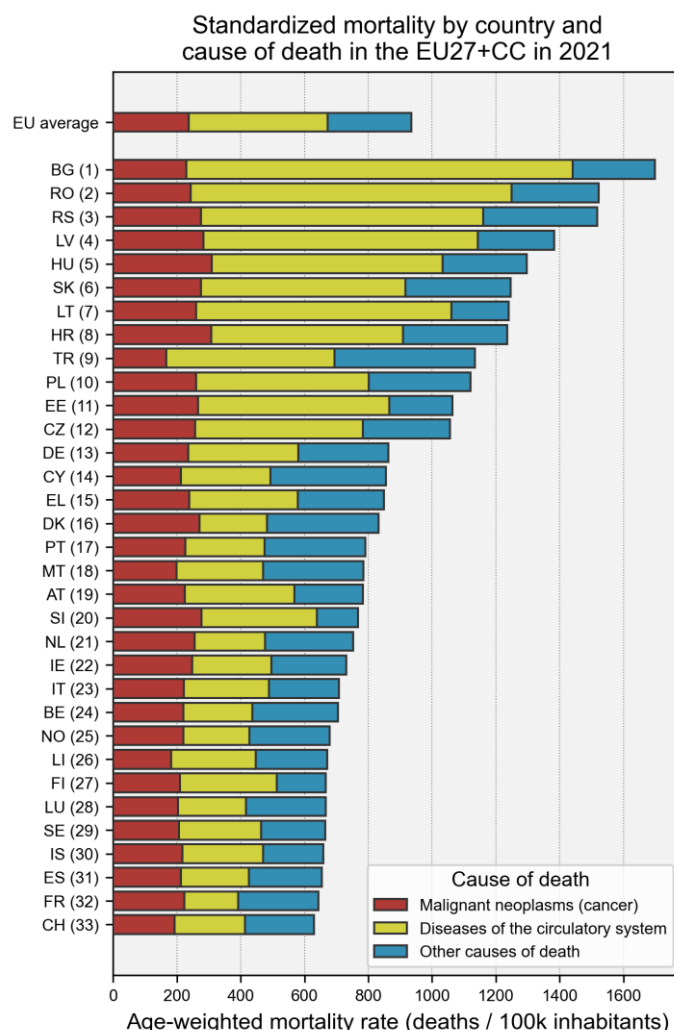


Figure 5-9. Stacked horizontal bar chart showing the 2021 national-level (NUTS0) overall age-weighted mortality and by cause of death in the EU27 and its Candidate Countries. A high-resolution version of this figure can be consulted [here](#).

5.3.3 Scenario 3 – Mapping two equal resolution indicators

5.3.3.1 Selected example data

With this third data availability scenario, we're entering the more challenging part of the chapter, wherein we will try to evaluate the interplay between multiple independent indicators (unlike the decomposed and hence dependent indicators shown in **Figure 5-6**) within a single map product. The outlining of a methodology to produce such synthetic visualisations, which may help to spatially unravel the cost and inequality in the health burden of pollution, is one of the main drivers of this report and a key contribution to the BEST-COST project.

Here we will combine example indicators from the previous two scenarios. We will visualise within a single map both the 2021 NO₂ air pollution exposure data, provided by EEA (pollution component) and EUROSTAT (exposure component), and the 2021 standardized mortality due to Diseases of the Circulatory System (DCS), which is abbreviated here to avoid overloading the map layout. Both indicators are to be mapped on NUTS2 level.



5.3.3.2 Mapping

When two or more variables are visualised within a single map display, we are applying multivariate cartography. A first example of multivariate cartography was already shown in the pie chart mapping of the previous data availability scenario (**Figure 5-6**). Pie chart mapping is only feasible when the mapped quantitative indicator can be decomposed in thematically related sub-indicators. When simultaneously mapping two independent variables, that may or may not be spatially correlated, **bivariate choropleth mapping (Text box 14)** could be a suitable approach.

Text box 14 - Bivariate choropleth mapping simultaneously displays two variables by combining colour gradients or patterns. Each colour or pattern represents a unique combination of the two variables' values, enabling the visualization of their spatial relationship. This method effectively highlights areas with specific combinations of characteristics, revealing correlations and interactions between the variables. Advantages include the ability to convey complex information in a single map and to identify spatial patterns that may not be apparent with separate maps. However, disadvantages include increased complexity, potential for misinterpretation, and difficulty in selecting colour schemes that remain distinguishable and intuitive for users.

Bivariate choropleth mapping is an interesting yet advanced form of cartography that presents challenges both on the production and interpretation side of the mapping. On the production side, the main challenges are firstly the limited or still experimental provision of freely available support tools for bivariate choropleth mapping within GIS software, and secondly the finding of a suitable bivariate colour scheme, a topic for which there are considerably fewer resources compared to univariate colour schemes. The bivariate choropleth maps made for this report were produced through combined use of QGIS, for the mapping itself, and a custom Python script, to produce a correct map legend. The interested reader is referred to the BEST-COST GitHub portal for more information. For the choice of the bivariate colour scheme, we drew inspiration from Joshua Stevens' blog post on this topic, [27] an often-cited source within the field of bivariate mapping.

On the interpretation side, we're aware that bivariate choropleth mapping is not yet well established within the research communities targeted by this report. As such, we may risk causing cognitive overload by just presenting such a map as it is, without preparation and context. For this reason, we've decided to first show the two univariate choropleth maps of NO₂ air pollution exposure (**Figure 5-10**) and standardised DCS mortality (**Figure 5-11**) to allow readers to predigest the spatial distributions of these two indicators within separate views. We use unipolar colour schemes for these univariate maps, with shades of red for NO₂ air pollution exposure and shades of blue for DCS mortality. Both maps have five value classes defined with quantiles classification (**Text box 5**) of the respective indicators. Missing values are in both cases denoted with a hatch pattern fill. Apart from the differing colour schemes, these maps are analogous to the examples discussed in the first data availability scenario.

After having shown the univariate maps, we then present the actual composed bivariate choropleth map in **Figure 5-12** that consolidates the former two. The used bivariate colour scheme displays shades of blue and red respectively for the classes having either high mortality or pollution exposure, and intermediate shades of purple for classes having lower or elevated values for both mortality and pollution exposure. The 2-dimensional nature of this



colour scheme is reflected by the colour grid legend, shown to the right of the map. To avoid visual overload, we use three quantile classes (**Text box 5**) per indicator, that together yield a total of nine classes. Note that the bivariate overlay of two quantiles classifications may yield empty combined classes. In this example however, each of the nine combined classes contain several NUTS2 regions.

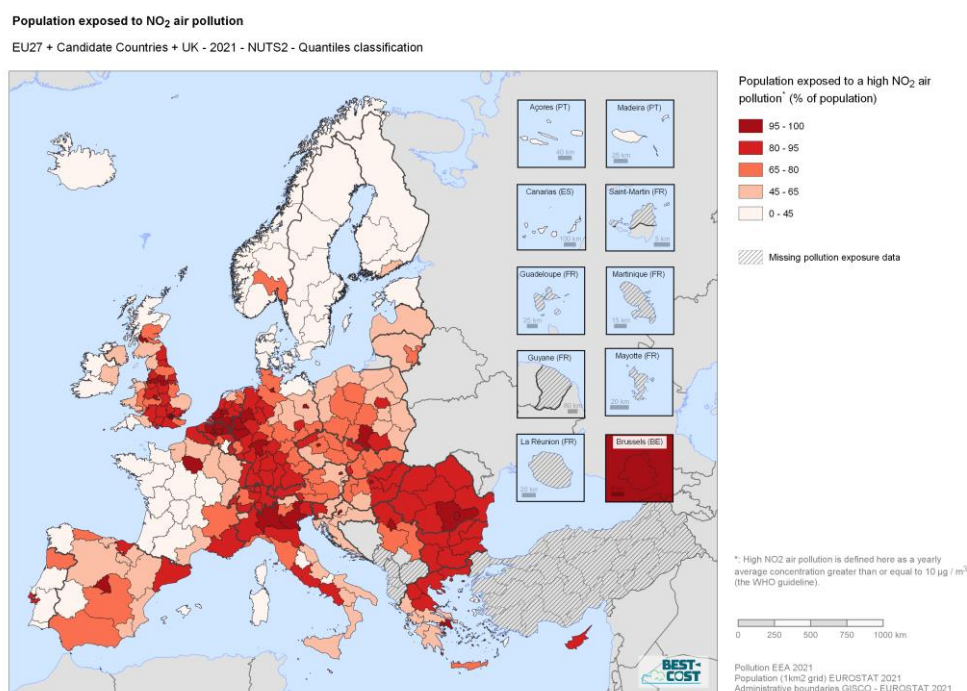


Figure 5-10. Choropleth with a unipolar colour scheme showing 2021 NO₂ air pollution exposure on NUTS2-level in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#).



Standardized mortality due to Diseases of the Circulatory System*
EU27 + Candidate Countries + UK - 2021 - NUTS2 - Quantiles classification

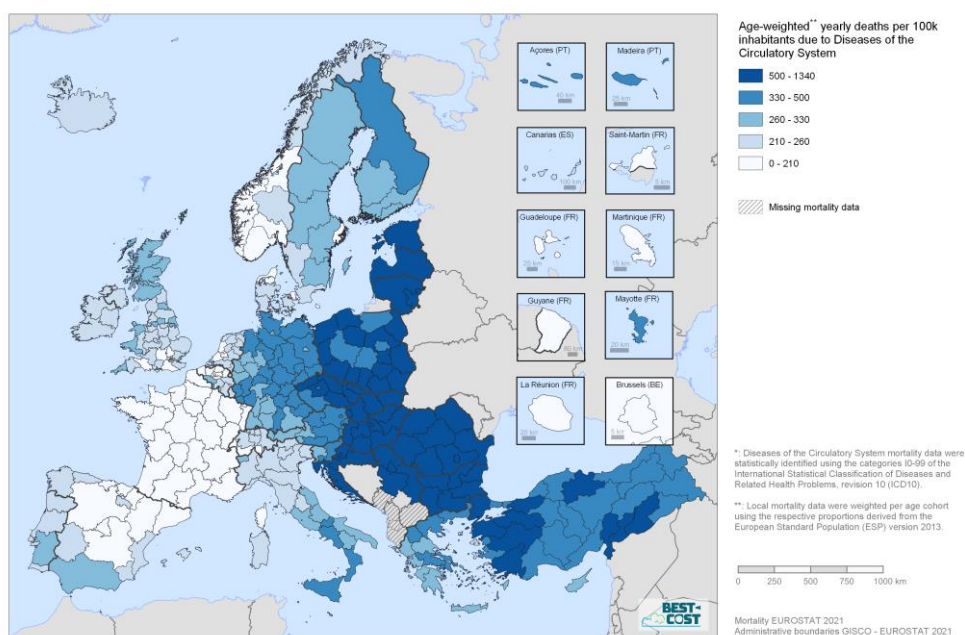


Figure 5-11. Choropleth map with a unipolar colour scheme showing 2021 age-weighted mortality due to diseases of the circulatory system on NUTS2-level in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#).



Standardized mortality due to Diseases of the Circulatory System* + population exposed to NO₂ air pollution

EU27 + Candidate Countries + UK - 2021 - NUTS2 - Quantiles classification (bivariate overlay)

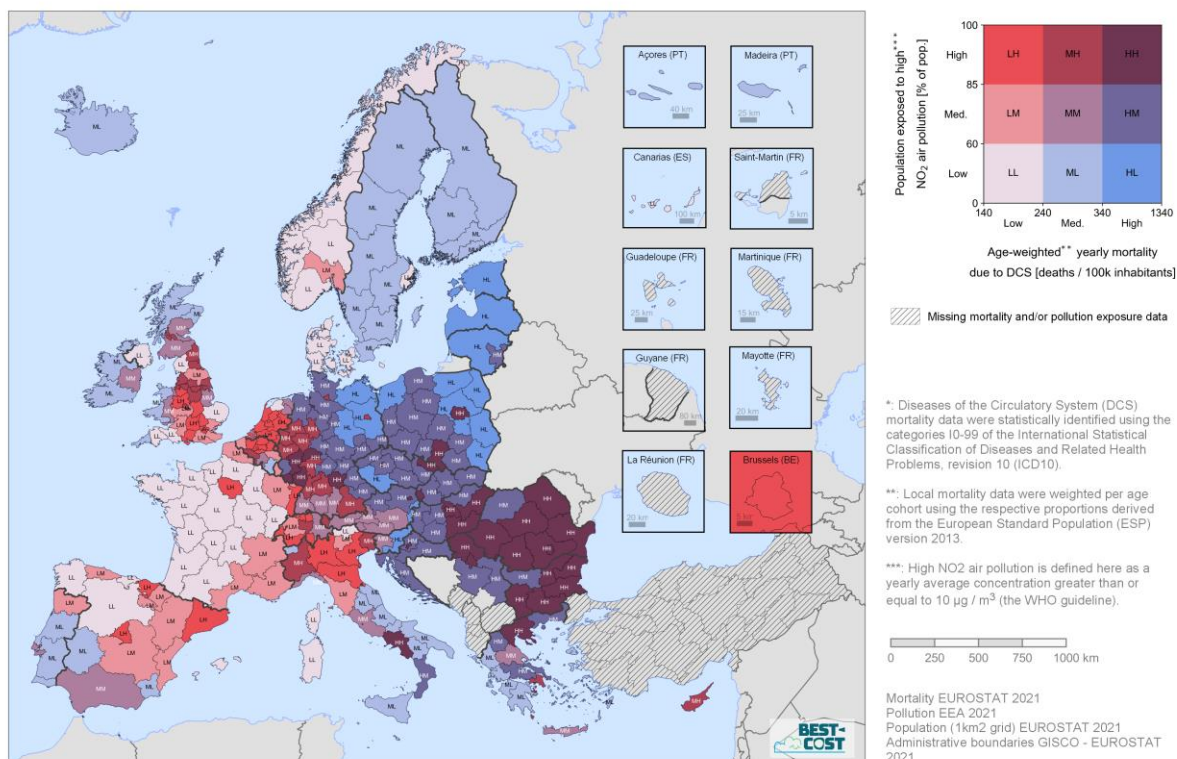


Figure 5-12. Bivariate choropleth map simultaneously showing 2021 NO₂ air pollution exposure and age-weighted mortality due to diseases of the circulatory system on NUTS2-level in the EU27, its Candidate Countries and the UK. The bivariate colour scheme is obtained by overlaying the unipolar colour schemes used in [Figure 5-10](#) and [Figure 5-11](#). A high-resolution version of this figure can be consulted [here](#).

On a more detailed design note, we propose to also plot combined class labels both within the NUTS2 regions on the map, if they aren't too small, as well as in the cells of the colour grid legend to help the reader interpret these classes using additional textual information. The first character of these two-letter labels indicates whether the mapped standardized DCS mortality value belongs to the low (L), moderate (M) or high (H) class, and the second character shows the same for NO₂ air pollution exposure. While the blue-red-purple bivariate colour scheme is considered colourblind-friendly, the labels do provide an additional readability safeguard. An inherent limitation of bivariate choropleth mapping, with regards to missing data, is that the mapping can't be performed for regions where data is missing for one of the two indicators.

5.3.3.3 Support graph

For the first support graph of this data availability scenario, we draw on **scatterplots** ([Text box 15](#)) to visualise the co-occurrence and co-distribution of the mapped continuous indicators on a 2D coordinate grid, with standardised DCS mortality on the horizontal axis and NO₂ air pollution exposure on the vertical axis ([Figure 5-13](#)). We also provide lateral histograms, plotted on the opposite sides of their respective axes, showing the data distributions of the separate indicators. The univariate colour coding used for the lateral histograms, and, by extension, the intermediate bivariate colour codes shown within the scatterplot match the colour schemes used for the bivariate choropleth map in [Figure 5-12](#)). These coloured



rectangles help the reader to make the conceptual link with the map and to assess the observation densities, i.e., the number of contained NUTS2 regions, for each combined bivariate class. NUTS2 regions are plotted with arrow symbols on the scatterplot. The colour and direction of the symbol indicates the EuroVoc sub-region [28] to which the NUTS2 region belongs, providing some geographic context to the plot.

Text box 15 - Scatterplots are graphical representations used to display the relationship between two continuous variables. Each point on the plot represents an observation, with its position determined by the values of the two variables on the x (horizontal) and y (vertical) axes. Scatterplots are useful for identifying correlations, trends, and outliers within datasets. They help in visualizing the direction and strength of relationships, as well as potential patterns or clusters. Scatterplots are widely used in statistical analysis, data exploration, and machine learning to understand and interpret complex relationships between variables.

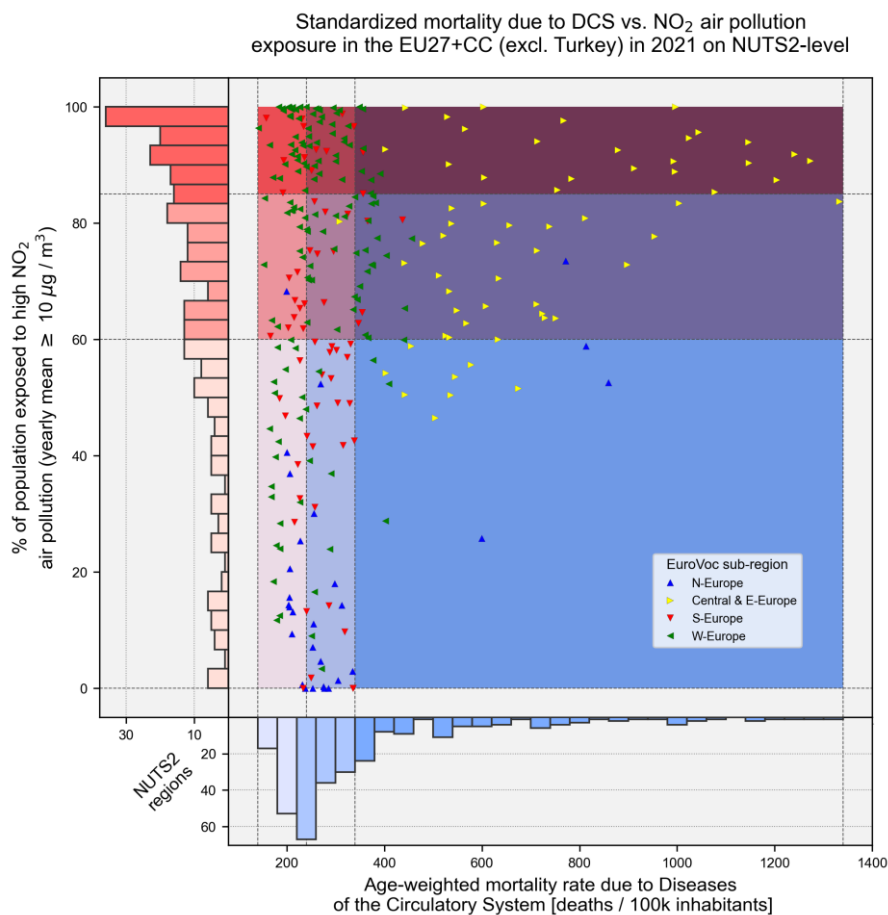


Figure 5-13. Colour coded scatterplot with lateral histograms showing the quantitative (co-)distributions of 2021 NO_2 air pollution exposure and age-weighted mortality due to diseases of the circulatory system on NUTS2-level in the EU27 (excluding Turkey) and its Candidate Countries. The colour coding of the scatterplot matches the bivariate colour scheme used in Figure 5-12. A high-resolution version of this figure can be consulted [here](#).

Next, in **Figure 5-14**, a bivariate elaboration is proposed of the bar chart plots with EU and national averages shown in the previous scenarios. Here, two bar charts are plotted in a vertically mirrored setup, on the left displaying the standardized DCS mortality averages and on the right the NO_2 air pollution exposure averages. To better organise this visualisation, the



rows are ranked in descending order on mortality, yielding an ordered ranking on the left and unordered ranking on the right. The rank indices are again added behind the two-letter country labels, and the colour coding of the bars matches the univariate schemes of the map.

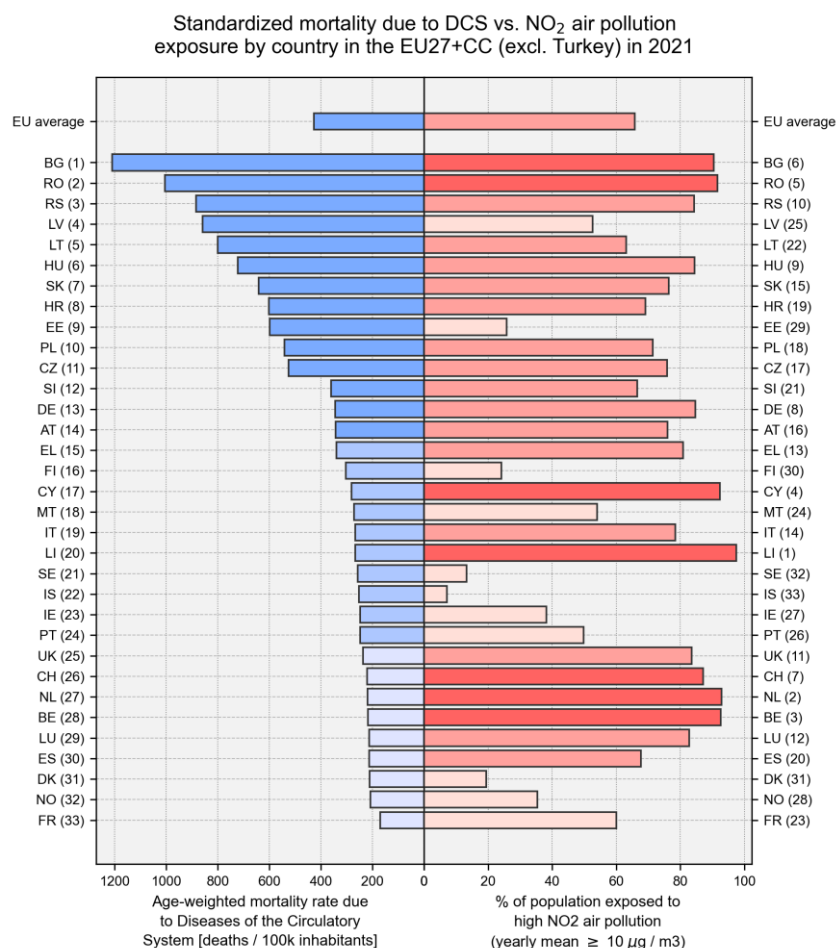


Figure 5-14. Mirrored horizontal bar chart showing 2021 national-level (NUTS0) NO₂ air pollution exposure, on the left, and age-weighted mortality due to diseases of the circulatory system, on the right, on NUTS2-level in the EU27 and its Candidate Countries (excluding Turkey). A high-resolution version of this figure can be consulted [here](#).

5.3.4 Scenario 4 – Mapping a higher and lower resolution indicator

5.3.4.1 Selected example data

For this data availability scenario, we'll be looking at the example of bivariate mapping of noise pollution exposure on NUTS3-level and standardized mortality due to DCS on NUTS1-level. The mortality indicator has already been introduced in the preceding scenarios, but it should be noted that temporal imputation was applied here to address missing NUTS1-level values. As was highlighted in the first chapter on data availability, noise pollution data is particularly problematic both on EU- and country-level. Modelled noise contours per transportation mode or activity, that form the basis for derived indicators, must according to the END Directive only be reported for major transport infrastructure and urban agglomerations exceeding a population of 100,000 inhabitants [29]. Even so, there are several countries that don't fully meet these reporting obligations. Consequently, we've decided to use an alternate noise indicator: the Quietness Suitability Index (QSI).



QSI is published by the EEA in raster format with a geographic resolution of 100m. So far, this layer has only been made available for the year 2016. QSI is a unitless indicator with relative significance, whose values range from 0 to 1. High QSI values indicate areas with high potential for being quiet [30]. Conversely, areas with low QSI values can be considered as having high potential for being noisy. As such, if we take the QSI Complement (QSIC), through the subtraction $1 - \text{QSI}$, we obtain a similarly bounded measure that rises with increasing noise potential. To obtain an indicator of noise pollution exposure, QSIC is first spatially aggregated to 1km^2 resolution and multiplied with the EUROSTAT population grid data. Then, the regional sums of the resulting measure are normalized using the corresponding total regional populations.

5.3.4.2 Mapping

Considering the diverging geographic resolutions, we chose a hybrid bivariate mapping approach for this scenario that overlays NUTS1-level mortality, visualised with line pattern mapping, on NUTS3-level noise exposure, displayed with a univariate choropleth base map (**Figure 5-15**). These mapping techniques have already been introduced separately in the first two scenarios covered in this chapter. Their combined use for this scenario plays into their respective strengths and the fact that they can be visually reconciled within a single map display (**Text box 8**, **Text box 13**).

For the noise exposure choropleth base map, we prefer a unipolar colour scheme (**Text box 9**), with shades of pink to distinguish it from previously mapped topics, as opposed to using a bipolar scheme which may risk overcharging the visual information density of the bivariate map. Five rounded quantile classes (**Text box 5**) are specified for the population weighted QSIC. Missing values are marked with dark grey fill rather than a hatch fill, to avoid confusion with the overlaid mortality map.

**Standardized mortality due to diseases of the circulatory system and population exposed to noise pollution**

EU27 + Candidate Countries + UK - 2016 (noise exposure) & 2021* (mortality) - NUTS1 (mortality) & NUTS3 (noise exposure) - Quantiles classification

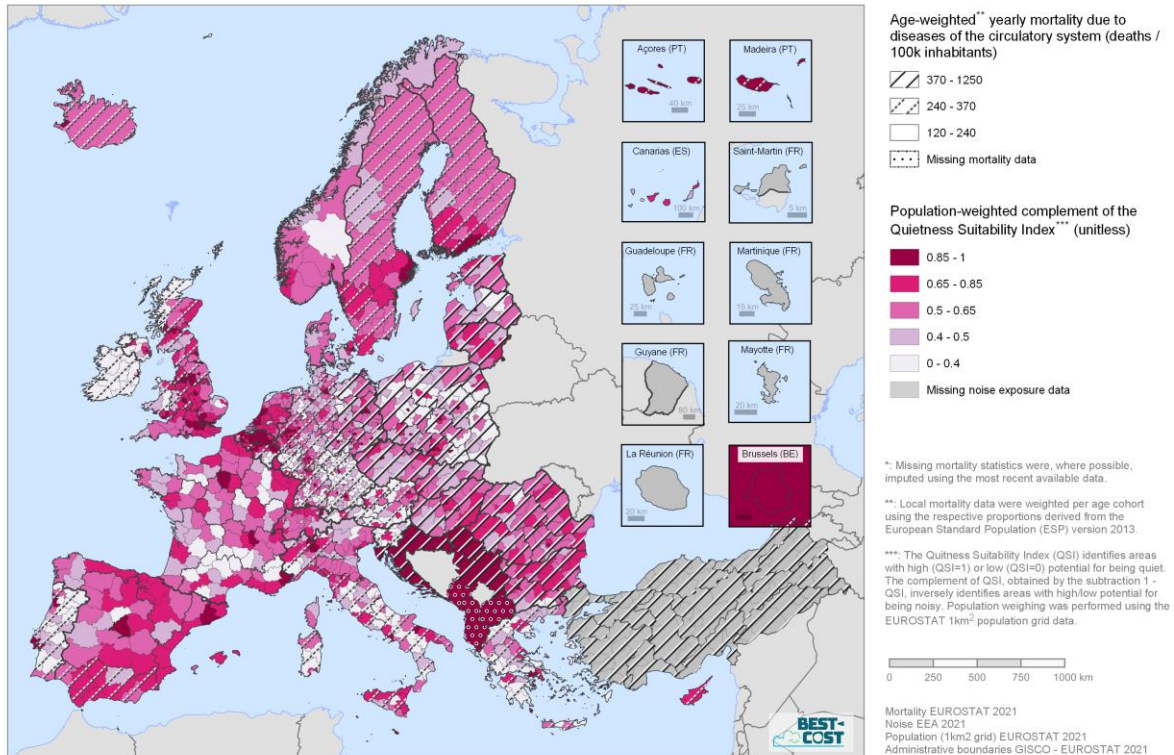


Figure 5-15. Combined line pattern map and choropleth map with a unipolar colour scheme, respectively showing 2021 age-weighted mortality due to diseases of the circulatory system, on NUTS1-level, and noise pollution exposure obtained with the Quietness Suitability Index, on NUTS3-level, in the EU27, its Candidate Countries and the UK. A high-resolution version of this figure can be consulted [here](#).

The line pattern mapping used to display standardized DCS mortality again draws on increasing line densities to denote higher mortality quantile classes, of which we use only three to avoid displaying too many different patterns. A design challenge of using this type of mapping here is making sure that the line patterns remain visible despite the varying colour and brightness of the underlying choropleth map, which is achieved by drawing both black and white lines in parallel. Missing mortality data are displayed with a point pattern fill, to distinguish them both from available mortality data as well as available or missing noise exposure data. Note that contrary to bivariate choropleth mapping, discussed in the previous scenario, it is possible here to handle missing values of the two mapped indicators separately. An additional design challenge of this map however lies in the display of three administrative levels, i.e., NUTS0 (countries), NUTS1 and NUTS3 regions, whose hierarchies are made discernible through border thickness and darkness.

5.3.4.3 Support graph

The support graph proposed for this data availability scenario, displayed in **Figure 5-16**, is a variation of the scatterplot with lateral histograms discussed in the previous chapter (**Figure 5-13**). Noise exposure is plotted on the vertical axis and standardized DCS mortality on the horizontal axis. The main design difference compared to the earlier discussed scatterplot lies



in the combined use of a unipolar colour scheme (**Text box 9**), for noise exposure, and line pattern fill, for

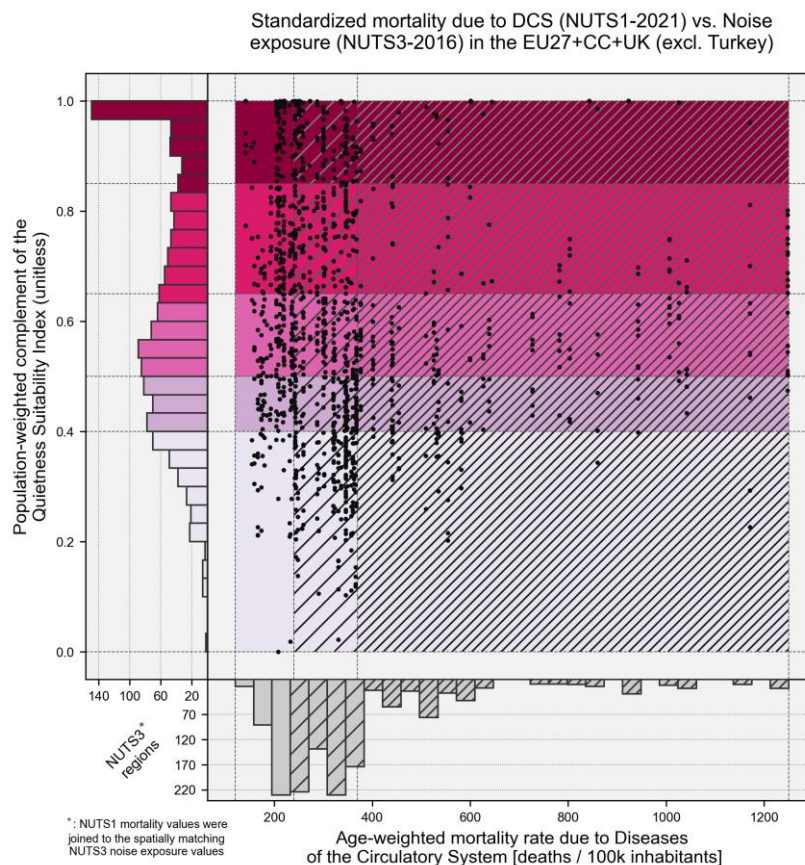


Figure 5-16. Colour coded and hatched scatterplot with lateral histograms showing the quantitative (co-)distribution of 2021 age-weighted mortality due to diseases of the circulatory system, on NUTS1-level, and 2016 population-weighted complement of the Quietness Suitability Index, expressing noise pollution exposure on NUTS3-level, in the EU27, its Candidate Countries (excluding Turkey) and the UK. The colour coding and hatch patterns of the different classes match those used in the map layout of **Figure 5-15**. A high-resolution version of this figure can be consulted [here](#).

Mortality. This approach yields rectangular regions in the scatterplot area of the plot that mimic the design of the corresponding bivariate map (**Figure 5-15**), and hence clarifies the conceptual link between the two. Because we inspect the co-distribution of a NUTS1- and NUTS3-level indicator in this plot, we must join NUTS3 noise exposure values geographically to NUTS1 mortality values. A drawback of this NUTS1-NUTS3 join, having a one-to-many **cardinality** (**Text box 16**), is a repetition of mortality values on the horizontal axis, an effect that can be observed in the graph through the vertically arranged point series. This geographic join is also mentioned in the graph itself, with a footnote on the bottom left.

Text box 16 – Cardinality in data management and database design refers to the uniqueness of data values in a column. High cardinality means a column has many unique values. Low cardinality means a column has a few unique values. Different types of cardinality include one-to-one, one-to-many, and many-to-many relationships. In one-to-one, each record in one table corresponds to one record in another. In one-to-many, one record in a table relates to multiple records in another. Many-to-many involves multiple records in both tables being related.



5.3.5 Scenario 5 – Mapping three equal resolution indicators

5.3.5.1 Selected example data

For this fifth and final data availability scenario, we will simultaneously assess three indicators on NUTS2-level. The following example data were chosen: 1. NO₂ air pollution exposure, 2. Standardized DCS mortality and 3. Socioeconomic deprivation. The first two of these indicators relate to the general topics of pollution exposure and health outcomes, that have already been covered by example data in the previous scenarios. The third topic of socioeconomic deprivation is only now introduced, seeing that within this work package of the BEST-COST project we are particularly interested in analysing social inequality within the health burden of environmental stressors. A three-way cartographic analysis of this kind requires some preparation and context, hence the decision to keep this for the last scenario.

A literature review of Multiple Deprivation Indices (MDI) was covered in report 3.1 of BEST-COST, and we refer the interested reader to this source for more information on the topic. To illustrate this scenario, we draw on the readily available Severe Material and Social Deprivation (SMSD) statistic made available by EUROSTAT within the frame of the EU Statistics on Income and Living Conditions (EU-SILC) framework. SMSD was chosen for practical purposes here, and it doesn't necessarily represent the most suitable MDI. SMSD expresses the percentage of a local population, down to NUTS2-level, living in conditions that meet at least seven out of thirteen deprivation criteria related to the individual and household [31]. Temporal imputation was used to find the most recent mortality, pollution exposure and deprivation values for each NUTS2 region, relative to the target year of 2021. Imputation by geographic aggregation (**Text box 4**) was also applied here, including for the UK, Ireland, and Estonia, to achieve a complete map at the cost of losing some spatial detail.

5.3.5.2 Mapping

The decision was made to draw on the relatively straightforward concept of **small multiples** (**Text box 17**) to implement the trivariate mapping of this scenario. We propose to start from bivariate choropleth mapping (**Text box 14**), that was already introduced in Scenario 3 to visualise the co-distribution of NO₂ air pollution exposure and standardized DCS mortality. Instead of trying to also consolidate SMSD within the same

Text box 17 - Small multiples in cartography involve presenting a series of similar maps, each showing different parts of a dataset or the same dataset over different time periods, arranged together or in close succession for easy comparison. This technique allows users to observe changes, trends, and patterns across multiple areas or timescales without switching contexts. Advantages include clarity in visual comparison, reduced cognitive load, and the ability to convey complex information. However, small multiples require sufficient space for effective display and may be challenging to interpret if too many maps are included or if the maps are too detailed.

singular view, we produce a series of four maps that each put into focus a subset of the mapped regions having a specific degree of deprivation. For this map series we define four quantile SMSD classes (**Figure 5-17**), that are respectively labelled very low, low, moderate, and high deprivation, and a map is made for each of these four deprivation classes (**Figure 5-18 - Figure 5-21**). The putting into focus of the NUTS2 region falling within the deprivation



class targeted by each map is achieved by blurring out the other NUTS2 regions, that fall outside the

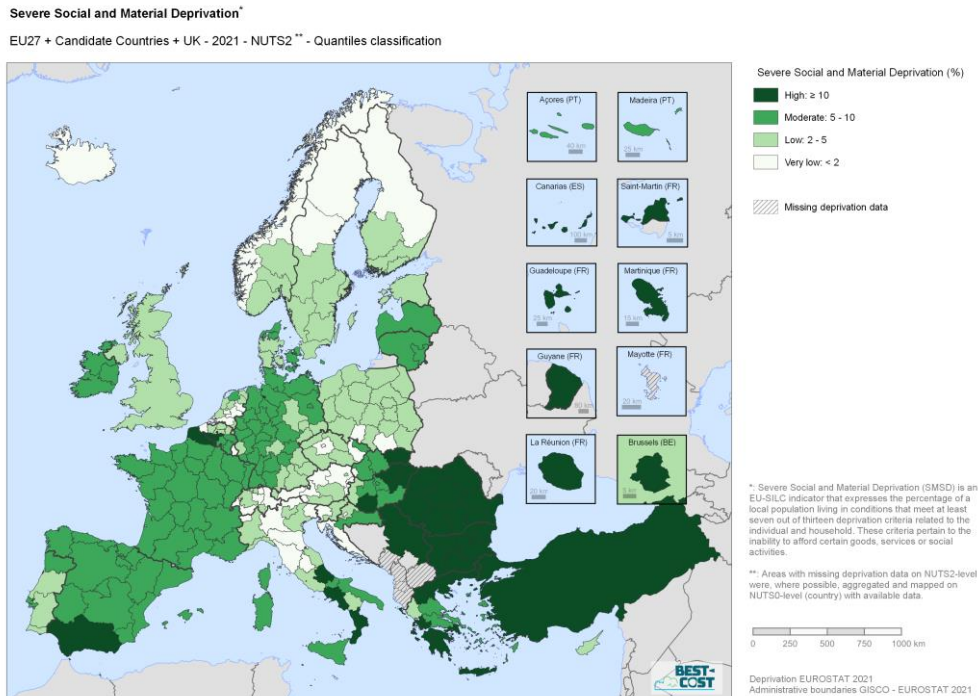


Figure 5-17. Choropleth map showing 2021 Severe Material and Social Deprivation on NUTS2-level in the EU27, its Candidate Countries and the UK. The four SMSD quantiles classes mapped here are re-used to define the focus areas of the subsequent maps **Figure 5-18 - Figure 5-21**. A high-resolution version of this figure can be consulted [here](#).

map scope. This blurring is implemented by overlaying the regions out of scope with a partially transparent and noisy point pattern fill. The targeted NUTS2 region will conversely have brighter and sharper colours, making them easy to spot and discriminate from the regions out of focus. The readability of the resulting maps will always partially depend on the underlying data, but it will be easier to observe spatial patterns if the NUTS2 regions belonging to a certain deprivation class and the underlying bivariate choropleth classes are somewhat clustered. For the sake of illustration, this is fortunately the case in the maps shown here.

Note that while small multiples are typically arranged within a single gridded view, we prefer to present the map series in close succession to better retain the spatial details in the separate maps.



Standardized mortality due to DCS^a vs. population exposed to NO₂ air pollution - Focus on very low deprivation areas (SMSD < 2%)

EU27 + Candidate Countries + UK - 2021^b - NUTS2^c - Quantiles classification (bivariate overlay)

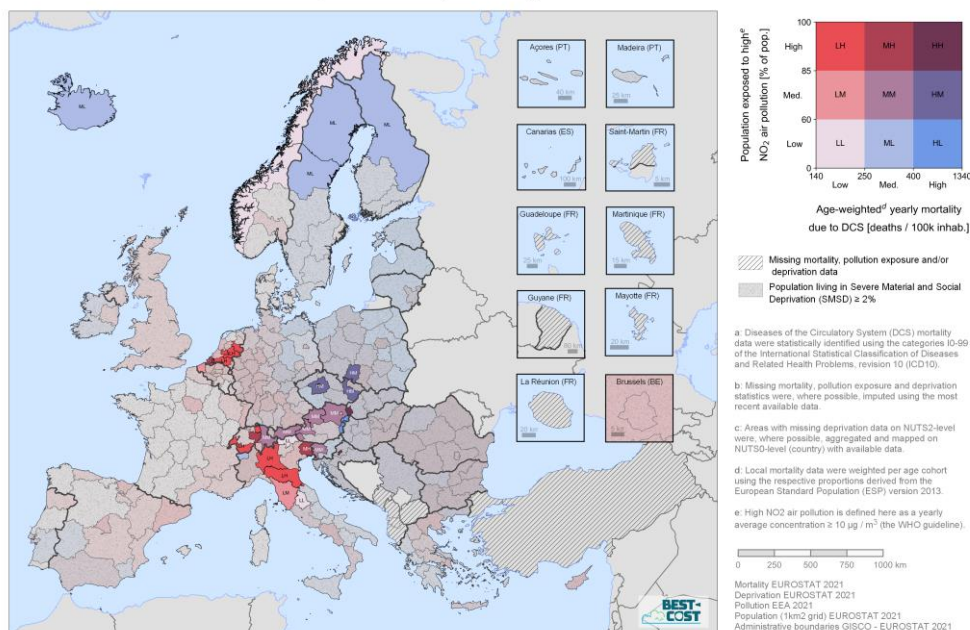


Figure 5-18. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a very low score (SMSD < 2%) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#).

Standardized mortality due to DCS^a vs. population exposed to NO₂ air pollution - Focus on low deprivation areas (2% ≤ SMSD < 5%)

EU27 + Candidate Countries + UK - 2021^b - NUTS2^c - Quantiles classification (bivariate overlay)

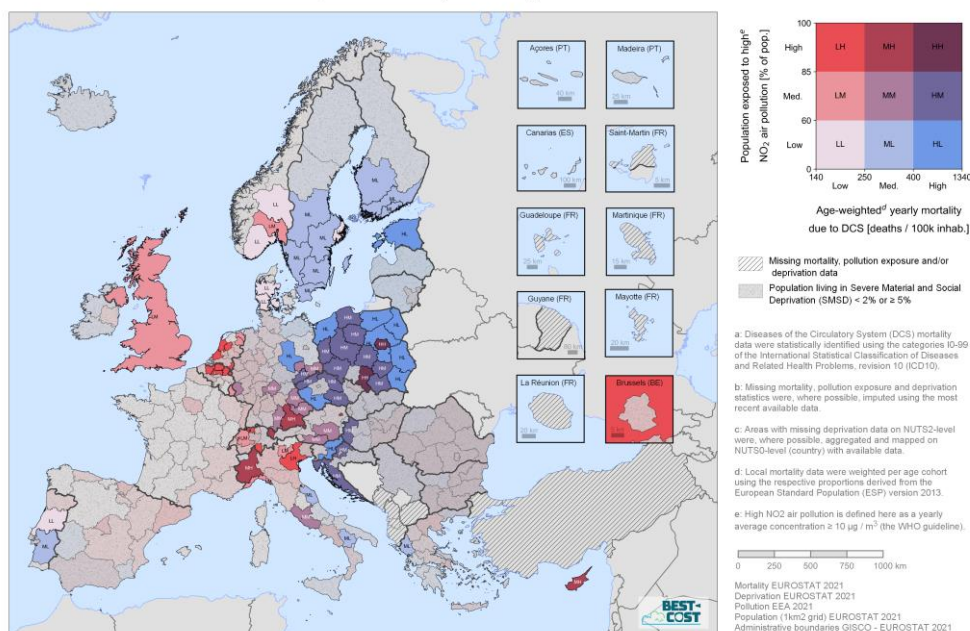


Figure 5-19. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a low score (2% ≤ SMSD < 5%) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#).



Standardized mortality due to DCS^a vs. population exposed to NO₂ air pollution - Focus on moderate deprivation areas ($5\% \leq \text{SMSD} < 10\%$)

EU27 + Candidate Countries + UK - 2021^b - NUTS2^c - Quantiles classification (bivariate overlay)

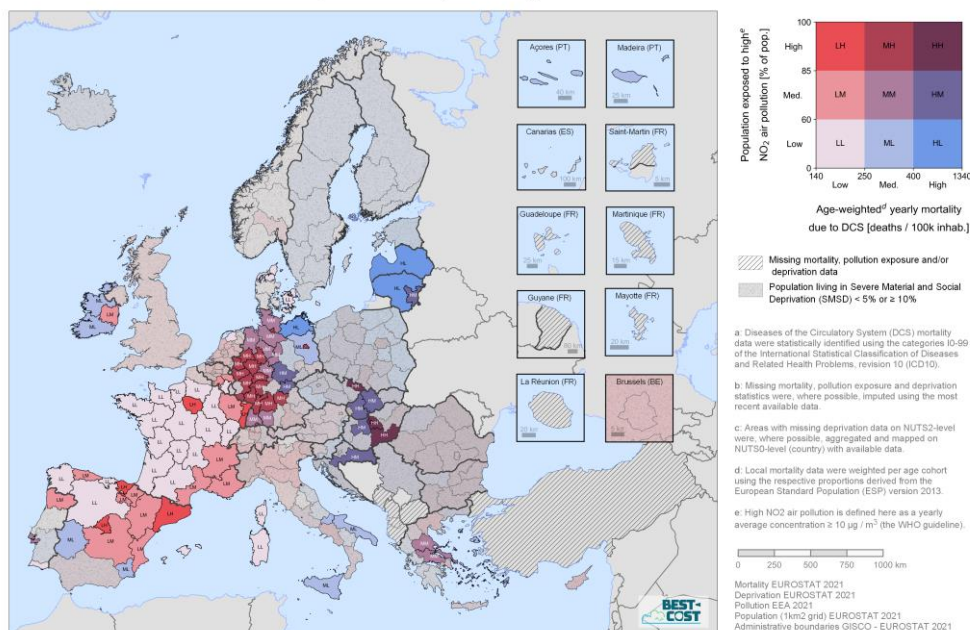


Figure 5-20. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a moderate score ($5\% \leq \text{SMSD} < 10\%$) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#).

Standardized mortality due to DCS^a vs. population exposed to NO₂ air pollution - Focus on high deprivation areas ($\text{SMSD} \geq 10\%$)

EU27 + Candidate Countries + UK - 2021^b - NUTS2^c - Quantiles classification (bivariate overlay)

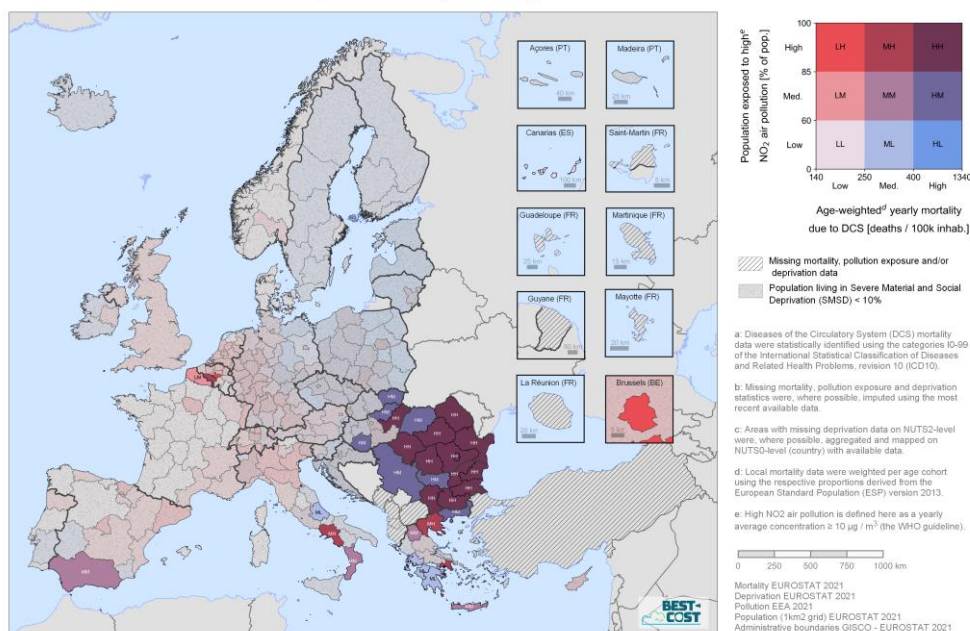


Figure 5-21. Bivariate choropleth map, shown and described earlier in **Figure 5-12**, now focusing on NUTS2 regions with a high score ($\text{SMSD} \geq 10\%$) for the Severe Material and Social Deprivation index. Regions out of scope are blurred out. A high-resolution version of this figure can be consulted [here](#).



5.3.5.3 Support graph

The support graph proposed for this trivariate data availability scenario is again an expansion on the scatterplot (**Text box 15**) with lateral histograms discussed earlier in Scenario 3 (**Figure 5-13**). The main design difference here lies in the shapes and colour coding of the symbols used to show the co-distribution of standardized DCS mortality (horizontal axis) and NO₂ air pollution exposure (vertical axis) in the scatterplot area of the graph. These shapes and colours are now adapted to reflect the SMSD classes targeted by each map. Shades of green are used to create visual contrast with the underlying bivariate red-blue colour scheme. We recognize that the blues and greens used in this support graph may be difficult to discern for people with colour vision deficiency, a design limitation that is mitigated by the brightness and shape differences of these symbols. Recall that in **Figure 5-13**, the scatterplot symbol shapes and colours were instead used to denote the EuroVoc sub-region corresponding to each observation.

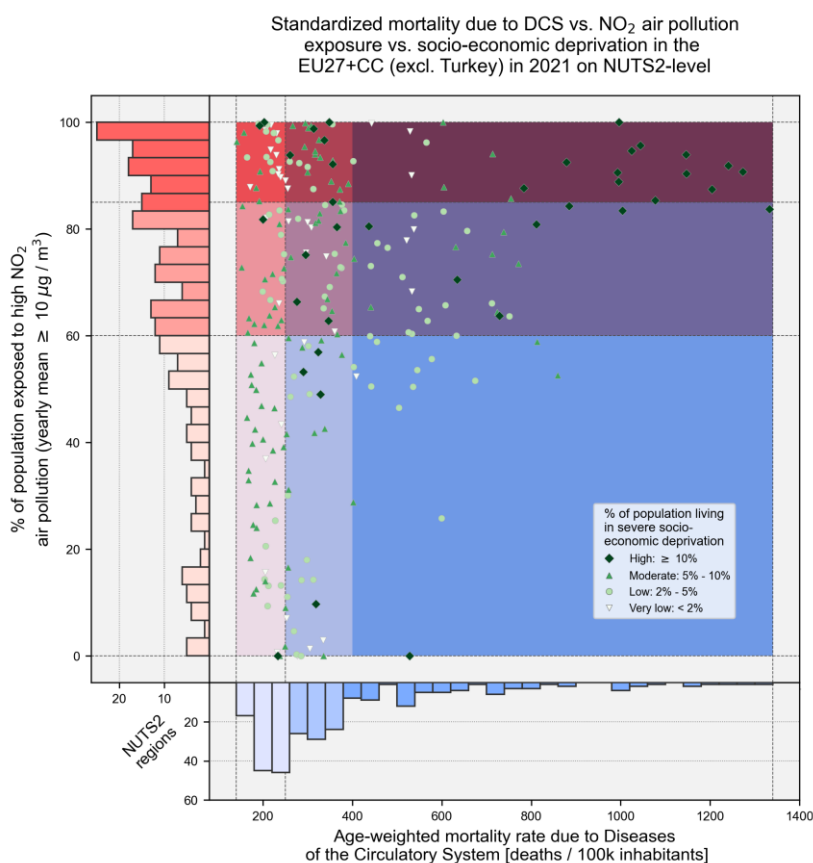


Figure 5-22. Colour coded scatterplot with lateral histograms, shown and described earlier in **Figure 5-13**, now having plot symbols whose shapes and colours reflect the varying degrees of Severe Material and Social Deprivation of the corresponding NUTS2 regions. A high-resolution version of this figure can be consulted [here](#).



5.4 Concluding remarks on the mapping

In this chapter, a methodological framework was put forward to assist researchers working on public health, environmental pollution and social inequality in visualizing various types and combinations of spatiotemporal statistics that may be of interest to them. The proposed mapping and graph designs provide a foundation for subsequent geographic and statistical analysis of the data pertaining to these topics, which can reveal patterns, interrelations between indicators, and other useful insights that can be difficult to obtain or communicate with alternate analytical techniques. Particular attention was paid to the challenge of the varying spatial resolutions on which these statistics are published, and on providing practical answers to multivariate mapping requirements. Clearly, it is not possible to address every possible situation or need that may arise, but the five scenarios put forward by the framework do already tackle a wide array of cartographic research challenges with matching mapping solutions. The scenarios of the framework are also presented in such an order so as to reflect a logical escalation of the depth of analysis, by first tackling univariate research questions, then bivariate and finally trivariate. We strongly recommend following this sequence whenever possible.

We would like to point out that additional and more intricate data availability scenarios were originally foreseen in this framework, including a scenario that addresses the mapping of three indicators on varying geographic resolutions (recall that the Scenario 5 does the same but on a single resolution). It became clear, however, after having completed the preceding scenarios, that such an exercise requires very elaborate mapping solutions that likely would, in the end, prove difficult to communicate to and be interpreted by policymakers and other stakeholders targeted by this work. This serves as a reminder that, while almost anything can be mapped, the investment required to produce such maps must be weighed against the lessons that can reasonably be learned from them by the intended audience.

The example maps discussed in each of the five scenarios of our framework present some interesting material on the interplay between pollution, health, and poverty, that may be tempting to study in more detail. However, we remind the reader that the maps presented in this chapter serve an illustrative purpose and are neither intended to be exhaustive nor finalized with regards to the specific needs of the BEST-COST consortium. For instance, health outcomes in this chapter were exclusively mapped using age-standardized mortality rates. This is useful and interesting information, yet by itself insufficient to assess the economic burden of disease, a key objective of the BEST-COST project. A more pertinent disease burden indicator catering to this need would be DALYs. Yet at the time of performing the preparatory work for this report, such data were not yet sufficiently available, neither on EU nor case country level, to produce clear example maps with. The mapping techniques proposed here remain valid however, regardless of the specific measures or study area used. We thus propose that the mapping, that was methodically outlined in this chapter, can be revisited at a later point in the project timeline, when more finalized datasets have been identified or produced.



6 Conclusion and recommendations

The exploration of the data availability of environmental exposures, health outcomes and socioeconomic deprivation was an initial step needed to undertake the estimation of the socioeconomic burden of environmental stressors in Europe. One of the key elements identified in this task was the variability in data availability and level of geographic resolution among the countries. This deviates from the initial ambition to be able to graphically depict and analyse burden estimates at high geographical resolution for the whole EU. For these reasons, the following recommendations should be taken into account when estimating the socioeconomic burden of environmental stressors in Europe:

1. Considering the different data availabilities, one needs to decide whether to focus on specific outcomes (e.g. mortality) or limit the analysis to one country in order to achieve a higher resolution. It is important to highlight that when using aggregated data some information will be lost. Particularly deprived and/or polluted areas might become less visible when using low geographical resolution data, which can hamper effective policy-making.
2. When comparing countries and/or EU areas, it is preferable to use low-resolution data to avoid losing information. According to our work, it seems feasible to calculate the socioeconomic burden of environmental stressors at NUTS 2 level.

Seeing the varying data availability and level of geographic detail among the EU and the selected case countries, we proposed a methodological framework that can be leveraged generically for the mapping of various health, environmental, and deprivation indicators. Five data availability scenarios are included in this framework to help researchers and analysts to produce clear meaningful maps on the socioeconomic burden of environmental stressors. The sequence of the five data availability scenarios reflects a systemic approach in which the depth of analysis is gradually increased to tackle increasingly complex research questions on the interrelations between these indicators, that may be mapped simultaneously on the same or different resolutions. The example maps and graphs included in this report serve as a reference and inspiration for reproducing this mapping on other cases. Essentially, we recommend that future mapping efforts performed within the BEST-COST project draw and elaborate on the guidelines and materials of the proposed framework. These mainly concern the following aspects:

3. We strongly recommend that the complete univariate-bivariate-trivariate sequence is followed when trying to tackle multivariate inquiries. Failing to do so may result in overlooking essential insights that can be gained from the data, and producing overloaded maps that are difficult to put into context. We also repeat that bivariate/trivariate mapping does not always yield clearly interpretable outputs, as the readability of the maps depend in part on the presence of underlying spatial patterns in the data.
4. The examples used in this report are meant to illustrate the methodological framework and are not intended to be final nor exhaustive with regards to the specific analytical needs of other BEST-COST work packages. As such, the mapping will need to be performed later in the project, particularly when more pertinent health outcome data,



in the form of DALY, are made available with a more extensive spatiotemporal coverage.



7 References

- [1] 'Global Burden of Disease and Risk Factors - Glossary. Lopez AD, Mathers CD, Ezzati M, et al., editors. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; New York: Oxford University Press; 2006.', The International Bank for Reconstruction and Development / The World Bank, 2006. Accessed: Sep. 09, 2024. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK11818/>
- [2] 'OECD. Health risks. Accessed on 24, June 2024. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/health-risks/indicator-group/english_1c4df204-en'.
- [3] 'WHO. Environmental risk factors and NCDs. Accessed on 12, September, 2024. Available on : <https://www.who.int/teams/noncommunicable-diseases/integrated-support/environmental-risk-factors-and-ncds>'. Accessed: Sep. 12, 2024. Available: <https://www.who.int/teams/noncommunicable-diseases/integrated-support/environmental-risk-factors-and-ncds>
- [4] 'Institute for Health Metrics and Evaluation (IHME). GBD Compare Data Visualization. Seattle, WA: IHME, University of Washington, 2024. Accessed Sep. 12, 2024. Available from <http://vizhub.healthdata.org/gbd-compare> (link is external).
- [5] 'GBD 2021 Risk Factors Collaborators. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet. 2024 May 18;403(10440):2162-2203. doi: 10.1016/S0140-6736(24)00933-4. PMID: 38762324; PMCID: PMC11120204.'
- [6] 'Hajat A, Hsia C, O'Neill MS. Socioeconomic Disparities and Air Pollution Exposure: a Global Review. Curr Environ Health Rep. 2015 Dec;2(4):440-50. doi: 10.1007/s40572-015-0069-5. PMID: 26381684; PMCID: PMC4626327.'
- [7] 'WHO Europe, 2012, Environmental health inequalities in Europe. Assessment report, World Health Organization, Regional Office for Europe, Copenhagen, Denmark'.
- [8] 'European Environment Agency. Unequal exposure and unequal impacts: social vulnerability to air pollution, noise and extreme temperatures in Europe, EEA Report No 22/2018. Copenhagen: EEA; 2019.'
- [9] 'Townsend, P. Journal of Social Policy , Volume 16 , Issue 2, April 1987 , pp. 125 - 146. DOI: <https://doi.org/10.1017/S0047279400020341>'.
- [10] 'ENVIRONMENTAL JUSTICE: Human Health and Environmental Inequalities. Robert J. Brulle and David N. Pellow. Annual Review of Public Health 2006 27:1, 103-124'.
- [11] 'European Environment Agency. Unequal exposure and unequal impacts: social vulnerability to air pollution, noise and extreme temperatures in Europe, EEA Report No 22/2018. Copenhagen: EEA; 2019.'
- [12] 'Nomenclature of territorial units for statistics (NUTS) - Eurostat. 2024. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Nomenclature_of_territorial_units_for_statistics_\(NUTS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Nomenclature_of_territorial_units_for_statistics_(NUTS)).' Accessed: Aug. 05, 2024. [Online]. Available: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Nomenclature_of_territorial_units_for_statistics_\(NUTS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Nomenclature_of_territorial_units_for_statistics_(NUTS))



- [13] 'Local administrative units (LAU) - Eurostat. 2024. <https://ec.europa.eu/eurostat/web/nuts/local-administrative-units>'. Accessed: Aug. 05, 2024. [Online]. Available: <https://ec.europa.eu/eurostat/web/nuts/local-administrative-units>
- [14] 'People at risk of poverty or social exclusion by NUTS 2 regions. Eurostat data browser. Last update : 08/04/2024 23:00. DOI : 10.2908/tgs00107. [Internet]. [cited 2024 Aug 05]. Available from: https://ec.europa.eu/eurostat/databrowser/view/tgs00107/default/table?lang=en&category=t_lilc.t_lilc_pe'.
- [15] 'Peterson, G.N. (2020). GIS Cartography: A Guide to Effective Map Design, Third Edition (3rd ed.). CRC Press. <https://doi.org/10.1201/9781003046325>'.
- [16] 'Slocum, T.A., McMaster, R.B., Kessler, F.C., & Howard, H.H. (2022). Thematic Cartography and Geovisualization (4th ed.). CRC Press. <https://doi.org/10.1201/9781003150527>'.
- [17] 'Countries - Eurostat. Accessed on Aug 14, 2024. <https://ec.europa.eu/eurostat/web/gisco/geodata/administrative-units/countries>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://ec.europa.eu/eurostat/web/gisco/geodata/administrative-units/countries>
- [18] 'Brewer, C. A., & Pickle, L. (2002). Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in Series. *Annals of the Association of American Geographers*, 92(4), 662–681. <https://doi.org/10.1111/1467-8306.00310>'.
- [19] B. Leff, A. Davis-Holland, D.-H. Consulting, E. Ducey, and B. A. Hamilton, 'BEST PRACTICES FOR MAP DESIGN - INTRODUCTION, 2016. https://proceedings.esri.com/library/userconf/fed16/papers/fed_86.pdf'.
- [20] 'NO2, European air quality data, (interpolated data). European Environment Agency. Accessed on Aug, 14, 2024. <https://www.eea.europa.eu/en/datahub/datahubitem-view/34bd02bf-be87-4122-b9a2-1e846c27a786>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://www.eea.europa.eu/en/datahub/datahubitem-view/34bd02bf-be87-4122-b9a2-1e846c27a786>
- [21] 'GEOSTAT - Eurostat. Accessed on Aug, 14, 2024. <https://ec.europa.eu/eurostat/web/gisco/geodata/population-distribution/geostat>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://ec.europa.eu/eurostat/web/gisco/geodata/population-distribution/geostat>
- [22] 'WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO.'
- [23] 'Directive - 2008/50 of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe - EN - EUR-Lex. <https://eur-lex.europa.eu/eli/dir/2008/50/oj>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://eur-lex.europa.eu/eli/dir/2008/50/oj>
- [24] 'Revision of the European Standard Population - Report of Eurostat's task force - 2013 edition. Accessed on Aug, 14, 2024. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-13-028>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-13-028>



- [25] 'Statistics | Eurostat. Causes of death - standardised death rate by NUTS 2 region of residence. Accessed Aug, 14, 2024. https://ec.europa.eu/eurostat/databrowser/product/page/HLTH_CD_ASDR2'. Accessed: Aug. 14, 2024. [Online]. Available: https://ec.europa.eu/eurostat/databrowser/product/page/HLTH_CD_ASDR2
- [26] 'ICD-10 Version:2019. WHO. Accessed Aug, 14, 2024. <https://icd.who.int/browse10/2019/en>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://icd.who.int/browse10/2019/en>
- [27] 'Bivariate Choropleth Maps: A How-to Guide. Joshua Stevens. Accessed Aug, 14, 2024. <https://www.joshuastevens.net/cartography/make-a-bivariate-choropleth-map/>'. Accessed: Aug. 14, 2024. [Online]. Available: <https://www.joshuastevens.net/cartography/make-a-bivariate-choropleth-map/>
- [28] 'EU Vocabularies - Publications Office of the EU. Accessed Aug, 14, 2024. <https://op.europa.eu/en/web/eu-vocabularies/concept-scheme/-/resource?uri=http://eurovoc.europa.eu/100277>', EU Vocabularies. Accessed: Aug. 14, 2024. [Online]. Available: <https://op.europa.eu/en/web/eu-vocabularies/concept-scheme/-/resource>
- [29] 'Environmental Noise Directive - European Commission. Accessed Aug, 14, 2024. https://environment.ec.europa.eu/topics/noise/environmental-noise-directive_en'. Accessed: Aug. 14, 2024. [Online]. Available: https://environment.ec.europa.eu/topics/noise/environmental-noise-directive_en
- [30] 'Quiet areas in Europe – The environment unaffected by noise pollution. European Environment Agency, 2016. <https://www.eea.europa.eu/publications/quiet-areas-in-europe>', European Environment Agency. Accessed: Aug. 14, 2024. [Online]. Available: <https://www.eea.europa.eu/publications/quiet-areas-in-europe>
- [31] 'Glossary:Severe material and social deprivation rate (SMSD). Eurostat. Accessed Aug, 14, 2024. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Severe_material_and_social_deprivation_rate_\(SMSD\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Severe_material_and_social_deprivation_rate_(SMSD))'. Accessed: Aug. 14, 2024. [Online]. Available: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Severe_material_and_social_deprivation_rate_\(SMSD\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Severe_material_and_social_deprivation_rate_(SMSD))



Appendix

Appendix 1 : data collection instruction form

Description of the extraction form

Page 1: Country information

Please indicate the country and the geographical areas commonly used to divide its territory (regions, counties, municipalities, other smaller statistical areas or census tracts). For each geographical area, indicate the corresponding European NUTS or LAU level, if applicable. If necessary, please give a small description of the geographical area (e.g. IRIS in France). Then indicate the number of entities in the administrative division, as well as the average, and the minimum and maximum range of the surface area and of the population.

NUTS: <https://ec.europa.eu/eurostat/web/nuts/background>

LAU: <https://ec.europa.eu/eurostat/web/nuts/local-administrative-units>

IRIS description (France): “All municipalities with more than 10,000 inhabitants, a part of municipalities with more than 5,000 inhabitant and less than 10,000; municipalities with less than 5,000 inhabitants are not divided and are an IRIS by themselves”

Page 2: Pollutant exposure

This page is divided into three parts:

1. Air pollutant: particulate matter (PM_{2.5}), ozone (O₃) and nitrogen dioxide (NO₂);
2. Noise pollutant: road traffic noise, railway noise and aircraft noise;
3. Population grid.

For each pollutant dispersion and population grid data source, please indicate:

- The metric and unit that expresses the level of pollution, or of the population (e.g., number or density).
- The geographical area for which exposure is available directly, or the resolution of the raster (e.g., cell size of the PM_{2.5} or population grid) or vector map (e.g., width of the noise contours) that will form the basis of the exposure assessment. You may



indicate **up to three sources**, the smallest areas or highest resolutions being the most important for the study.

- The completeness of the exposure numbers (e.g., nation-wide, regional...) or the geographical territory that is covered by the pollution map or population grid.
- The reference period (e.g., year). If data are available for multiple periods, indicate the most recent.
- The type of the data (example: meteorological station, administrative data, model output, satellite image).
- The accessibility of the data, by choosing one of four options:
 - o The data is owned by the institution of the user and can be used freely.
 - o The data is owned by another institution, but is available as open access.
 - o The data is owned by another institution, and can be used given authorisation of the owner (e.g., by signing a contract).
 - o The data is owned by another institution, and cannot be used.
- The source of the data (institution in charge) and a weblink to the data in case of open access.
- Whether exposure or population numbers are stratified by age and/or sex (not applicable for raster or vector data) and if so, please specify the strata (e.g., definition of age groups).
- Whether socio-economic indicators accompany the exposure or population numbers and if so, please specify which ones (not applicable for raster or vector data).

You can add any additional information that may be useful in the last column.

Page 3: Disease/disturbance prevalence

On this page, the **diseases and noise disturbance/consequences** (high sleep disturbance, high annoyance, cognitive impairment) relevant for the study are listed. For stroke, depending on the data available in your country: please give information on each of the level 4 strokes available (ischemic stroke, intracerebral hemorrhage and subarachnoid hemorrhage), otherwise provide information of the level 3 stroke.

Following the disease models developed in task 1.1 (WP1), diseases are defined according to ICD codes (the International Classification of diseases (ICD) - 9th and 10th revisions - see table below). Noise annoyance and sleep disturbance do not have any ICD code.

For each health outcome, please indicate :



- The geographical area for which the measurement is available. You can indicate **up to three geographical areas**, the smallest being the most important for the study.
- The reference period (e.g., year). If data are available for multiple periods, indicate the most recent.
- The definition used for the disease during the measurement of prevalence (example: ICD code, self-reported, medication consumption, other type of codes)
- The data input used (example: hospital discharge, national registries, survey, administrative data)
- The accessibility of the data, by choosing one of four options:
 - o The data is owned by the institution of the user and can be used freely.
 - o The data is owned by another institution, but is available as open access.
 - o The data is owned by another institution, and can be used given authorisation of the owner (e.g., by signing a contract).
 - o The data is owned by another institution, and cannot be used.
- The source of the data (institution in charge) and a weblink to the data in case of open access.
- Whether exposure numbers are stratified by age and/or sex and if so, please specify the strata (e.g., male vs female, age groups).
- If socio-economic indicators were included during pollutant measurement, please specify which ones

You can add any additional information that may be useful in the last column.

Page 4: Disease mortality

Within this page, please provide the same above-described information for the **disease-related mortality**.

Diseases with their ICD codes

Disease level 3 (according to GBD)	Disease level 4 (according to GBD)	ICD code
Lung cancer		ICD9: 162–162.9, 209.21, V10.1-V10.20, V16.1-V16.2, V16.4-V16.40 ICD10: C33, C34–C34.92, Z12.2, Z80.1-Z80.2, Z85.1-Z85.20
Ischemic Heart Disease		ICD9: 410–414 ICD10: I20–I25
	Myocardial infarction	ICD9: 411 ICD10: I21–I23



	Angina Pectoris	ICD9: 413 ICD10: I20
Stroke		ICD10: G45–G46.8, I60–I63.9, I65–I66.9, I67.0–I67.3, I67.5–I67.6, I68.1–I68.2, I69.0–I69.3
	Ischemic stroke	G45–G46.8, I63–I63.9, I65–I66.9, I67.2–I67.3, I67.5–I67.6, I69.3
	Intracerebral hemorrhage	I61–I62, I62.1–I62.9, I68.1–I68.2, I69.1–I69.2
	Subarachnoid hemorrhage	I60–I60.9, I62.0, I67.0–I67.1, I69.0
Type II Diabetes		E11.2, E11.21, E11.22, E11.29
Chronic Obstructive Pulmonary Disease		ICD 9: 491–492, and 496 ICD 10: J41, J42, J43, J44, and J47
Asthma		ICD9: 493 ICD10: J45 and J46
Heart failure, unspecified		ICD10: I50. 9
Disorders of newborn related to slow fetal growth and fetal malnutrition	Newborn small for gestational age	ICD10: P05.1
Essential (primary) hypertension		ICD10: I10
Overweight, obesity and other hyperalimentation	Obesity	ICD10: E66.0
Disorders of bone density and structure	Osteoporosis	ICD10: M80, M81, M82
Chronic kidney disease		ICD10: N18
Disturbance of activity and attention		ICD10: F90.0
Depressive episode		ICD10: F32
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue (C00–C75)	Malignant neoplasm of bladder	ICD10: C67
	Malignant neoplasm of kidney, except renal pelvis	ICD10: C64
	Skin cancer (non-melanoma)	ICD10: C44

Appendix 2 : excel data collection file (country information, pollution exposure, diseases and disturbance prevalence and diseases mortality - example : Belgium)

Appendix 2.1 : country information (demographics and geographic areas)

Country	Geographical area	Corresponding NUTS or LAU	Description of area (if necessary)	Number of units (if available)	Population			Surface area (km²)		
					Mean (if available)	Min (if available)	Max (if available)	Mean (if available)	Min (if available)	Max (if available)
Belgium	Region	NUTS 1		3	3872541	1228655	6709787	10222,33	162	16906
	Province	NUTS 2		11	1038896,8	294400	1890627	2787,91	162	4461
	Arrondissement	NUTS 3/ LAU 1		44	268339,2	50896	1228655	696,98	162	1597
	Municipality	LAU 2		589	20254,4	77	543165	2,90	1,2	215
	Statistical sector	/	Census tracts	19795	569	0	8569	1,32	0,01	44,83

Appendix 2.2 : pollutant exposure (air and noise)

Air pollutant	Metric and unit	Geographical area/ resolution	Geographical coverage/	Reference period	Data type	Data accessibility	Data source (+ URL)	Stratification information	If yes which indicator	Additional information
PM2.5	Annual mean concentration in µg/m³	10 m	Belgium	2021	Gridded model output (ATMO-Street)	The data is owned by another institution, but is available as open access.	IRCEL - CELINE*	NA	NA	
O3	Annual mean of daily maximum 8-hour mean concentration in µg/m³	10 m	Belgium	2021	Gridded model output (RIO-IFDM)	The data is owned by another institution, but is available as open access.	IRCEL - CELINE*	NA	NA	
NO2	Annual mean concentration in µg/m³	10 m	Belgium	2021	Gridded model output (ATMO-Street)	The data is owned by another institution, but is available as open access.	IRCEL - CELINE*	NA	NA	

IRCEL-CELINE* url : <http://ftp.irceline.be/atmostreet/>

EEA** url : <https://www.eea.europa.eu/en/datahub/datahubitem-view/c952f520-8d71-42c9-b74c-b7eb002f939b>



Noise pollutant	Metric and unit	Geographical area/ resolution	Geographical coverage/	Reference period	Data type	Data accessibility	Data source (+ URL)	Stratification information	If yes which indicator	Additional information
Road traffic noise	Lden in dB	Brussels Capital Region	Brussels Capital Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Brussels Capital Region	Brussels Capital Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lden in dB	Flemish Region	Flemish Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Flemish Region	Flemish Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lden in dB	Walloon Region	Walloon Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Walloon Region	Walloon Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
Railway noise	Lden in dB	Brussels Capital Region	Brussels Capital Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Brussels Capital Region	Brussels Capital Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lden in dB	Flemish Region	Flemish Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Flemish Region	Flemish Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lden in dB	Walloon Region	Walloon Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Walloon Region	Walloon Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
Aircraft noise	Lden in dB	Brussels Capital Region	Brussels Capital Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Brussels Capital Region	Brussels Capital Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lden in dB	Flemish Region	Flemish Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	
	Lnight in dB	Flemish Region	Flemish Region	2016	Exposure number per 5 dB band	The data is owned by another institution, but is available as open access.	EEA**	No	NA	



Appendix 2.3 : diseases and disturbance prevalence

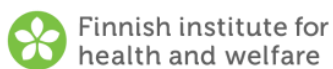
Disease level 3	Disease level 4	Geographical area	Reference period	Definition of disease	Incidence/prevalence	Data type	Data accessibility	Data source (+ URL)	Stratification information available	If yes which indicator	Additional information
Lung cancer		Region	2021	ICD-10	incidence	registry	The data is owned by a	Belgian cancer reg	Yes	Age and sex	
		Region	2020	ICD-10	10-year prevalence	simulated data from incidence	The data is owned by t	Sciensano	Yes	Age and sex	
Ischemic Heart Disease		Region	2020	ICD-10	incidence	hospital discharge data	The data is owned by a	Technical minimum	Yes	Sex	
Stroke	Unspecified*	Region	2020	ICD-10	incidence	hospital discharge data	The data is owned by a	Technical minimum	Yes	Sex	It is called "cere"
		Province	2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex and SES	
	Ischemic stroke	-	-	-	-	-	-	-	-	-	-
	Intracerebral hemorrhage	-	-	-	-	-	-	-	-	-	-
	Subarachnoid hemorrhage	-	-	-	-	-	-	-	-	-	-
Type II Diabetes		Region	2020	ATC codes (co	prevalence	health reimboursment da	The data is owned by t	Sciensano	Yes	Age and sex	no distinction be
		Province	2020	ATC codes (co	prevalence	health reimboursment da	The data is owned by a	IMA	Yes	Age and sex	no distinction be
		Province	2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex and SES	no distinction be
Chronic Obstructive Pulmonary		Province	2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex and SES	Includes Chronic
Asthma		Province	2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex and SES	
Heart failure, unspecified		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
Sleep disturbance		Province	2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex and SES	different types c
Annoyance		Province	2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex and SES	different types c
Cognitive impairment		-	-	-	-	-	-	-	-	-	-
Disorders of newborn related to slow fetal growth and fetal	Newborn small for gestational age	Region	2022		prevalence	register	The data is owned by a	Studiecentrum voc	Yes	Age, parity, weight,...	
Essential (primary) hypertension		Region	2018	Self-reported	prevalence	examination survey	The data is owned by t	Sciensano	Yes	Age, sex, education	
eight, obesity and other hyperalimen	Obesity	Region	1997-2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano - https://	Yes	Age, sex, SES	
sorders of bone density and structur	Osteoporosis	Region	1997-2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano	Yes	Age, sex, education	
Chronic kidney disease		Region	2013-2021	ICD-10	incidence		The data is owned by t	Sciensano	Yes	Age, sex	
Disturbance of activity and attention		-	-	-	-	-	-	-	-	-	-
Depressive episode		Region	2001-2018	Self-reported	prevalence	health interview survey	The data is owned by t	Sciensano - https://	Yes	Age, sex, SES	
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue	Malignant neoplasm of	Region	2004-2022	ICD-10	incidence	register		Belgian Cancer Reg	Yes	Age, sex	
	Malignant neoplasm of kidney, except renal pelvis	Region	2004-2022	ICD-10	incidence	register		Belgian Cancer Reg	Yes	Age, sex	
	Skin cancer (non-melanoma)	Region	2004-2022	ICD-10	incidence	register		Belgian Cancer Reg	Yes	Age, sex	



Appendix 2.4 : diseases mortality

Disease level 3	Disease level 4	Geographical area	Reference period	Definition of disease	Data type	Data accessibility	Data source (+ URL)	Stratification information available	If yes which indicator	Additional information
Lung cancer		Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	NA	
		region	2020	ICD-10	mortality registry	The data is owned by the institution	sciensano	Yes	Age and sex	
Ischemic Heart Disease		Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
		region	2020	ICD-10	mortality registry	The data is owned by the institution	sciensano	Yes	Age and sex	
Stroke	Unspecified*	region	2020	ICD-10	mortality registry	The data is owned by the institution	sciensano	Yes	Age and sex	It is called "cerebro"
	Ischemic stroke	Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
	Intracerebral hemorrhage	Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
	Subarachnoid hemorrhage	Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
Type II Diabetes		Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
		region	2020	ICD-10	mortality registry	The data is owned by the institution	sciensano	No	Age and sex	possibility to distinguish
Chronic Obstructive Pulmonary Disease		Statistical sector		ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
		region	2020	ICD-10	mortality registry	The data is owned by the institution	sciensano	Yes	Age and sex	
Asthma		Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
		region	2020	ICD-10	mortality registry	The data is owned by the institution	sciensano	Yes	Age and sex	
Heart failure, unspecified		Statistical sector	2020	ICD-10	mortality registry	The data is owned by another institution	statbel	No	no	
Disorders of newborn related to slow fetal growth and fetal malnutrition	Newborn small for gestational age	-	-	-	-	-	-	-	-	-
Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue	Malignant neoplasm of bladder	Country	2021	ICD-10		The data is owned by another institution	Fondation con	No		
	Malignant neoplasm of kidney, except renal pelvis	Country	2021	ICD-10		The data is owned by another institution	Fondation con	No		
	Skin cancer (non-melanoma)	Country	2021	ICD-10		The data is owned by another institution	Fondation con	No		

Partners





www.best-cost.eu



@BestCostEU



Funded by
the European Union

BEST-COST is funded by the European Union's Horizon Europe programme under Grant Agreement No.101095408.