Calculation and Validation of a New Synthetic and Autocorrelate Index to Measure the Health Status of a Population

Domenico Tebala

Italian National Institute of Statistics

Giovanni Domenico Tebala (✉ giovanni.tebala@ouh.nhs.uk)

Oxford University Hospitals NHS Foundation Trust   https://orcid.org/0000-0001-7152-4096

Research article

Abstract

Background

The status of health of an individual and - more broadly - of a community or population is defined by the status of their determinants of health. A "systemic" approach to define the health determinants is necessary in order to explore the complex relations existing among them. This study is aimed at identifying a 'composite systemic' index of health to measure the impact of socioeconomic factors on public health at local level and to analyze possible spatial autocorrelations between neighboring regions.

Methods

A Composite Index of Health (CIH) was constructed on the basis of known indicators of socio-economic well being by using the COMIC (COMposite Indices Creator) Software and was validated on the Italian population and a nationwide comparison has been performed.

Results

Analysis of the determinants showed a significant direct correlation between health, environment, work and wealth and inverse correlation between health and social distress. The analysis of data from Italian provinces confirmed the South-North gradient of well-being.

Conclusions

The CIH is a reliable and robust index to evaluate the health of a local population. Although it was validated on Italian data, the index can be easily adapted to any Country.

Background

The status of health of an individual and - more broadly - of a community or population is defined by the status of their determinants of health. Their unequal distribution is a well-known problem of public health [1,2]. Despite clear recommendations on how to best address these inequalities, there is still a lack of clarity on how to target political interventions. This can be due, at least in part, to the imperfect knowledge of what data should be considered as health indicators and, within the known factors, the role that socio-economic indicators play in the determination of the general status of health of a population. Several Authors have highlighted the obvious strict causal relation existing between socioeconomic factors and status of health, where inequalities of the first would lead to disparities in the latter [3]. The study of single determinants may not be considered completely reliable, so public health researchers are turning their attention towards a

complex systems approach at local and national level, which can explain interlinked social and health inequalities [4,5]. Such a systemic approach may allow us to study the different components of the complex health system and to get to the bottom of their intricate relationships, while obtaining the wide picture on how the system – as a whole – works [4]. This would also prompt public health professionals and politicians to plan the right interventions to tackle disparities and inequalities at local and national level.

The present study is aimed at identifying a 'composite index of health' (CIH) to measure the impact of socioeconomic factors on public health at provincial level through the ESWB (Equitable and Sustainable Well-Being) methodology and to analyze possible spatial autocorrelations between different regional areas.

In fact, data measured in a specific area (or province) can be influenced by what happens in nearby areas, generating what is commonly called "spatial autocorrelation" or "spatial interdependence". In this regard, the LISA indicators (Local Indicator of Spatial Association) provide a local framework to the measure of autocorrelation, enabling each spatial unit (i.e., the province) to assess the degree of spatial association and similarity with the surrounding elements.

In the case of positive autocorrelation, these associations can be of the High-High type (high values observed in a territorial unit and high values also in its vicinity) or Low-Low type (low values observed in a territorial unit and low values also in its neighboring areas). Conversely, in the case of negative self-correlation, the associations will be of the High-Low or Low-High type. In all other cases, there will be no autocorrelation or non-significant autocorrelation.

Methods

For simplicity, the CIH has been constructed on the basis of the Italian provinces but can be easily adapted to any other Country.

The method of construction of the index followed these steps:

1) analysis of the theoretical framework, methodology used and indicators;

2) choice of the statistical methodology

3) statistical analysis: in order to assess the robustness of the identified method and, therefore, improve decision-making, we also completed an influence analysis to analyze the most significant indicators (software COMIC - COMposite Indices Creator by Italian National Institute of Statistics – ISTAT)

4) analysis of the results with the help of a georeferenced map of the CIH of Italian provinces and a Cluster Map LISA which shows the provinces with statistically significant values of the LISA index,

classified in five categories: a) Not significant (white); b) High-High (red); c) Low-Low (blue); d) Low-High (light blue); e) High-Low (light red) (software GeoDa v.1.12).

Process to calculate the indicator

To evaluate the status of health of a community, it is necessary to identify and analyze the various determinants that affect health and, when negative, the lack of health. It is a complex set of factors relative to the individual and to their role in the society: personal factors such as gender and age, personal behavior and lifestyle, social factors, living conditions, work, access to health services, financial status and environment.

The approach used involves the construction of macro areas (pillars) by aggregating elementary indicators. Both pillars and elementary indicators have been considered non-replaceable. To construct the synthetic CIH, we adopted the indicators and polarity considered in the 2015 document "Fair and sustainable wellbeing in towns" [7] published by ISTAT and defined in Table 1.

Methodology

The matrix of data on Italian provinces was divided into four progressive steps:

a) Selection of a set of basic indicators on the basis of an ad hoc evaluation model centred on the existence of quality requirements;

b) Further selection aimed at balancing the set of indicators within the theoretical framework of the structure. Outcome indicators are impact indicators as the ultimate result of an action of a stakeholder's activity or process;

c) Calculation of synthetic indices (pillars), by making use of the more appropriate methodology to obtain usable analytical information on the health of Italian provinces;

d) Processing of a final synthetic index as a rapid empirical reference concerning the degree of health of Italian provinces.

Missing values were attributed via the hot-deck imputation and, where not possible, were considered overlapping Italy's average value.

The choice of the method of synthesis is based on the assumption of a formative measurement model, in which it is believed that the elementary indicators are not replaceable, which is to say, cannot compensate each other.

The exploratory analysis of input data was performed by calculating mean, average standard deviation and frequency, as well as building up a correlation matrix and performing a principal component analysis. Since this is a non-compensatory approach, the simple aggregation of elementary indicators

was carried out using the correct arithmetic average with a penalty proportional to the "horizontal" variability.

Normalization of primary indicators took place by conversion into relative indexes compared to the variation range (min-max).

Attribution of weights to each elementary indicator has followed a subjective approach, opting for the same weight for each of them. Since, in some cases, the elementary indicators showed different polarity, it was necessary to reverse the sign of negative polarities by linear transformation.

Calculation of the CIH has been performed with the Adjusted Mazziotta-Pareto Index (AMPI), which is used for the min-max standardization of elementary indicators, and aggregated with the mathematical average penalized by the "horizontal" variability of the indicators themselves. In practice, the compensatory effect of the arithmetic mean (average effect) is corrected by adding a factor to the average (penalty coefficient) which depends on the variability of the normalized values of each unit (horizontal variability), or by the variability of the indicators compared to the values of reference used for the normalization.

The synthetic index of the *i-th* unit, which varies between 70 and 130, is obtained by applying, with negative penalty, the correct version of the penalty method for variation coefficient (AMPI +/-), where:
$$AMPI_i^- = M_{ri} - S_{ri} c v_i.$$

M_{ri} e S_{ri} are, respectively, the arithmetic mean and the standard deviation of the normalized values of the indicators of the *i* unit, and $cv_i = S_{ri} / M_{ri}$ is the coefficient of variation of the normalized values of the indicators of the *i* unit.

The correction factor is a direct function of the variation coefficient of the normalized values of the indicators for each unit and, having the same arithmetic mean, it is possible to penalize units that have an increased imbalance between the indicators, pushing down the index value (the lower the index value, the lower the level of health).

This method satisfies all requirements for the wellbeing synthesis:

- Spatial and temporal comparison
- Irreplaceability of elementary indicators
- Simplicity and transparency of computation
- Immediate use and interpretation of the obtained results
- Strength of the obtained results

An influence analysis was also performed to assess the robustness of the method and to verify if and with which intensity the composite index rankings change following elimination from the starting set of a primary indicator. This process allowed us also to analyze the most significant indicators.

The analysis was conducted with the COMIC software, that allows the calculation of synthetic indices and rankings, as well as the comparison of different synthetic methods to select the most suitable among them and write an effective report based upon results.

Approval by an Ethical Committee was not necessary as this study did not involve human participants but only widely available population data. For this reason, the principles of the Declaration of Helsinki [8] do not apply to this study.

Results

Table 2 shows significant correlations between the analyzed indicators of the macroareas, especially the "direct" correlations between health and environment (r = 0.81), work (r = 0.61) and material well-being (r = 0,57) and the "reverse" correlation between health and social distress (r =-0,48). The influence analysis describes the indicators that most influence the composition of rosters in health of Italian provinces. In Figure 1 it is evident how variables of work (s = 10.96) and health (s = 9.15) are the most significant.

The cartographic representation of the final CIH value, alongside with the descriptive analysis of data, yields the usual dualistic pattern South vs Center-North of Italy as in other domains of ESWB and shows how the health of the Italian population is closely related to the socio-economic components.

The best CIH performances (Figure 2) are grouped in Tuscany (Florence, Prato and Pisa) and Lombardy (Lecco, Monza and Como), but the "healthiest" province is Trento (CIH = 111.36 - CIH for Italy = 100), thanks mainly to the cultural (123.58) and employment (117.30) indexes which in that region (Trentino) "weigh" more than the more strictly sanitary variables. Frosinone (Lazio) occupies the last position in the ranking list (CIH = 83.22), even though it is the whole South to be heavily penalized (55% of the provinces below the average belong to the South of Italy).

Our results show a good spatial interaction between neighbouring provinces (Moran's index = 0.39) (Figure 3). This is particularly true in the South, where is a definite positive autocorrelation between 12 provinces, with low values observed in a province and low values in the surrounding areas (Low-Low Blue Color) (Figure 4).

Discussion And Conclusions

The 'composite index of health' (CIH) combines statistical rigor with high level of communicability and is useful to facilitate comparison and analysis of the status of health of a population at local and national level. However, in-depth studies need the analysis of single indicators.

The evident North-South CIH gradient is a further demonstration of the well-known social, cultural and financial disparities within the Country. Despite being quite autonomous social and political entities, neighbouring provinces often share the same cultural background in particular in the South. It is therefore

possible to identify spatial patterns able to describe areas of "multidirectional dependence", where contiguous areas show similar levels of the same phenomenon.

Improving the general status of health of the population would entail tackling the disparities by means of "upstream" changes directed to change the social-environmental determinants of health rather than health itself [3], although an improvement of local healthcare system would also act as a driver to increase the social wellbeing and would back up the improvement of the other indicators.

The governance framework to the global determinants of health should have a "systemic" and conjoined approach involving all the stakeholders and all the national and local policies [3]. Improving the determinants of health would be of great advantage not only to the individuals but also, we may say "mostly", to the society in terms of improved productivity, increased tax revenue and lower welfare and healthcare costs [9]. The UK Marmot Report clearly identified six objectives of the political choices: (a) improving the health of children, (b) maximizing each person capabilities, at every age, (c) creating good work environments, (d) ensuring healthy lifestyles, (e) improving the status of the environment, (f) creating policies and procedure for medical prevention.

Clearly, reaching these objectives would require a thorough knowledge of the determinants of health within the society and within each layer of it.

The CIH would be useful to have an idea of the general status of health as determined by social and financial factors and to guide government actions on non-health issues which may yield positive health outcomes.

In conclusion, the CIH is a good indicator of the status of health of a population at local and national level and may be extremely useful to explore the relations between health and socioeconomic factors and also to evaluate the influence of proximity with high or low-performing regions on the health of citizens of a specific region or province. Although hereby validated on the Italian population, the CIH can be easily generalised to any Country.

Abbreviations

CIH
Composite Index of Health
COMIC
COMposite Indices Creator
ESWB
Equitable and Sustainable Well-Being
LISA
Local Indicator of Spatial Association
ISTAT
National Italian Institute of Statistics

AMPI
Adjusted Mazziotta-Pareto Index

Declarations

Ethics approval and consent to participate: Approval by an Ethical Committee was not necessary as this study did not involve human participants but only widely available population data. For this reason, the principles of the Declaration of Helsinki [8] do not apply to this study. Furthermore, as this study did not involve any patient, informed consent was not necessary.

Consent for publication: not applicable

Competing interests: the Authors declare that they have no competing interest to disclose.

Funding: no funding was available for this manuscript

Authors' contributions: Both DT and GT conceived and planned this study; DT retrieved and analysed the data; GT reviewed the analysis of data; both DT and GT discussed the data, drafted the manuscript and approved its final version. Both DM and GT are equally accountable for this paper.

Acknowledgements: none

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding Author on reasonable request, however, most of the data used are freely available online.

References

1. Commission on Social Determinants of Health. Closing the gap in a generation. Geneva: WHO, 2008. Available from:
http://www.who.int/social_determinants/final_report/csdh_finalreport_2008_execsumm.pdf

2. Organisation for Economic Cooperation and Development. Rising inequality: youth and poor fall further behind. Insights from the OECD Income Distribution Database. Paris: OECD, 2014. Available from: https://www.oecd.org/social/OECD2014-Income-Inequality-Update.pdf

3. Carey G, Crammond B. Systems change for the social determinants of health. BMC Public Health 2015;16:662.

4. Baum F, Lawless A, Williams C. Health in All Policies from International Ideas to Local Implementation: Policies, Systems and Organizations. In: Clavier C, de Leeuw E, Eds. Health promotion and the Policy Process. London: Oxford University Press, 2013. p.188-217

5. Fisher M, Milos D, baum F, Friel S. Social determinants in an Australian urban region: a complexity lens. Health Promotion International [Internet], 2014 Aug 8, Available at
heapro.oxfordjournals.org/cgi/doi/10.1093/heapro/dau071

6. De Savigny D, Adam T, Alliance for Health Policy and Systems Research, World Health Organization. Systems thinking for health systems strengthening [Internet]. Geneva: Alliance for Health Policy and Systems Research: World Health Organization; 2009. Available from: http://public.eblib.com/choice/publicfullrecord.aspx?p=476146

7. UrBes 2015 - Il benessere equo e sostenibile nelle città. Available at https://www.istat.it/it/files//2015/04/UrBes_2015.pdf. Istat, 2015. Last access 26 May 2020

8. WMA Declaration of Helsinki. Ethical principles or medial research involving human subjects. World Medical Association, 9th July 2018. Available at https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/. Last access 28th May 2020

9. Marmot M, Goldblatt P, Allen J et al. Fair society healthy lives (The Marmot Review). Strategic review for health inequalities in England post-2010. Available at http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review/fair-society-healthy-lives-full-report-pdf.pdf . Last access 25 May 2020

Tables

Table 1: Indicators and polarity

cro areas	Indicators	Polarity
vironment	Surface intended for urban gardens (square meters per 100 inhabitants)	+
	Total density of the green areas (protected natural areas and urban green areas) (percentage on municipal surface)	+
	Green historic density and urban parks of significant public interest (m2 to 100 m2 of built-up areas)	+
	Maximum number of days exceeding the limit for the protection of human health provided for PM10 in the provincial capitals	-
	Dispersions of the drinking water network (% values)	-
	Noise Control in which at least one has been detected exceeding the limits in the provincial capitals (per 100,000 inhabitants)	-
	Density of bike paths in the provincial capitals - (km per 100 km2)	+
	Availability of pedestrian areas in the provincial capitals (m2 per 100 inhabitants)	+
cial tress	Rate of homicides reported by police to the court for the province (per 100,000 inhabitants)	-
	Thefts in homes reported by the police to the court (per 100,000 inhabitants)	
	Robberies reported by the police to the court (per 100,000 inhabitants)	
	Rate per province road accidents (per 100,000 inhabitants)	
lture	Number of public libraries (per 100,000 inhabitants)	+
	Number of museums, archaeological sites and monuments (per 100,000 inhabitants)	
	Visitors of non-state institutes of antiquities and art institute for (number per 1,000 visitors)	
	Number of users of public libraries (per 100 inhabitants)	
alth	Life expectancy at birth by sex and province (average number of years)	+
	Infant mortality rate by sex and province (rates per 10,000 live births)	-
	Standardized mortality rate for transport accidents for people aged 15-34 by sex and province (rates per 10,000 residents)	-
	Standardized mortality rate from cancer for people 20-64 years old by gender and province (rates per 10,000 residents)	-
	Standardized mortality rate for dementia and nervous system diseases for people aged 65 and over by gender and province (rates per 10,000 residents)	-

rk	Employment rate of the population aged 20-64 years by province	+
terial ll-being	Per capita GDP	+

Table 2: Correlation matrix of the macro areas

Macro areas	Environment	Social desease	Culture	Health	Work	Material well-being
Environment	1,00					
Social desease	-0,03	1,00				
Culture	0,33	0,28	1,00			
Health	0,81	-0,48	0,32	1,00		
Work	0,50	0,01	0,53	0,61	1,00	
Material well-being	0,57	-0,08	0,09	0,54	0,24	1,00

Figures
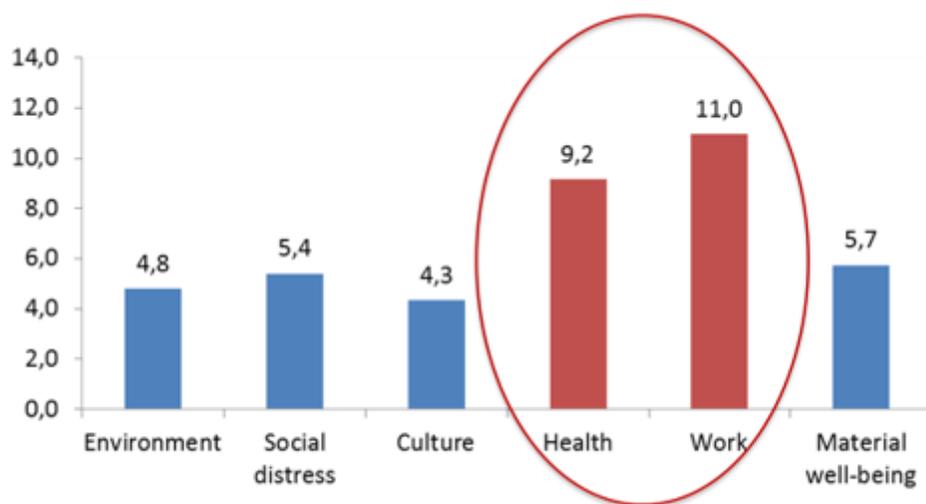
Figure 1

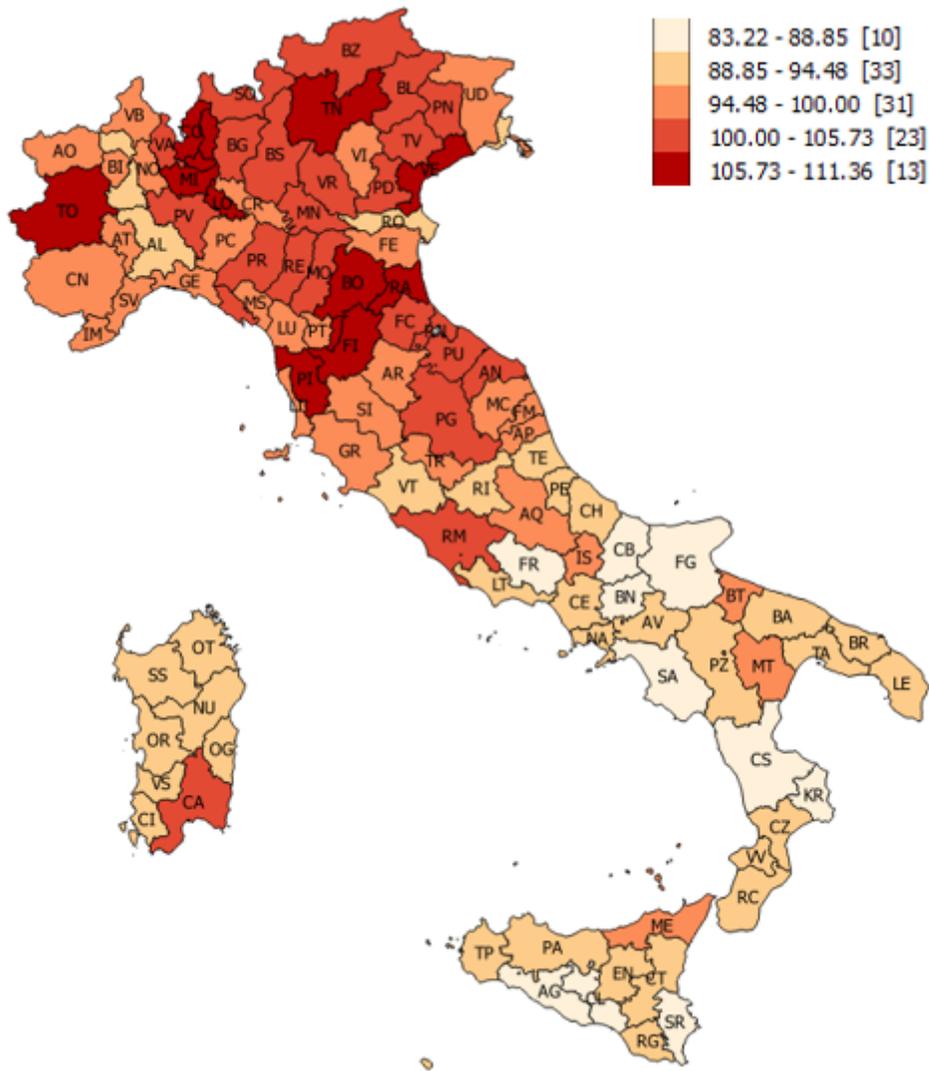Influence Analysis: of the shifts for basis indicator of macro areas

Figure 2

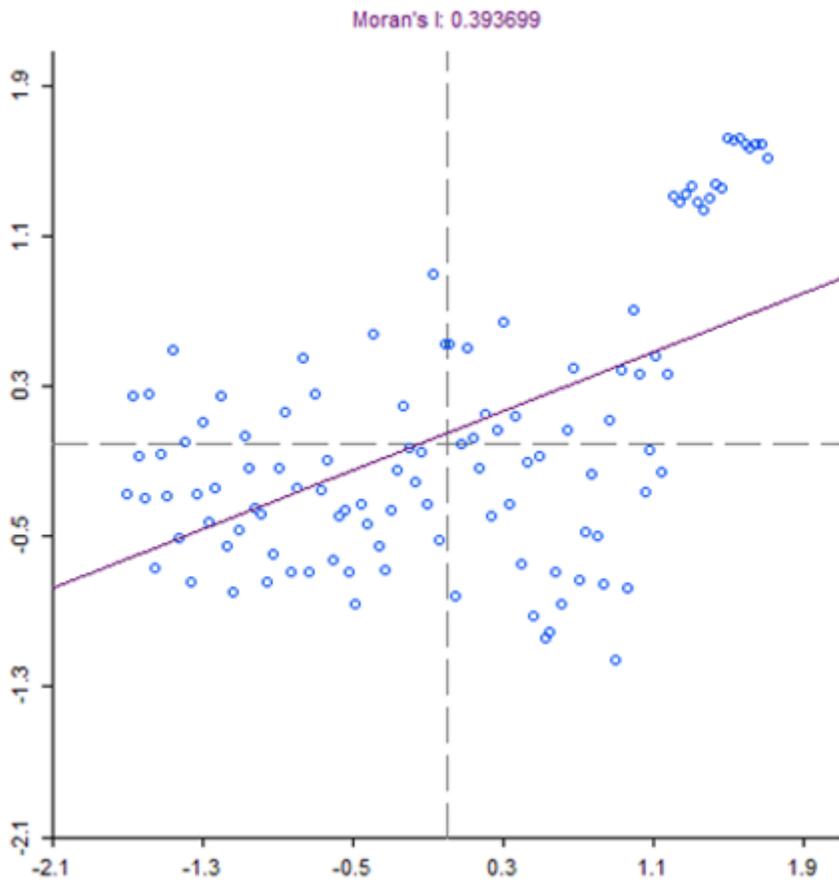Territorial distribution of the Composite Index of Health. (In square brackets the number of provinces in that range)
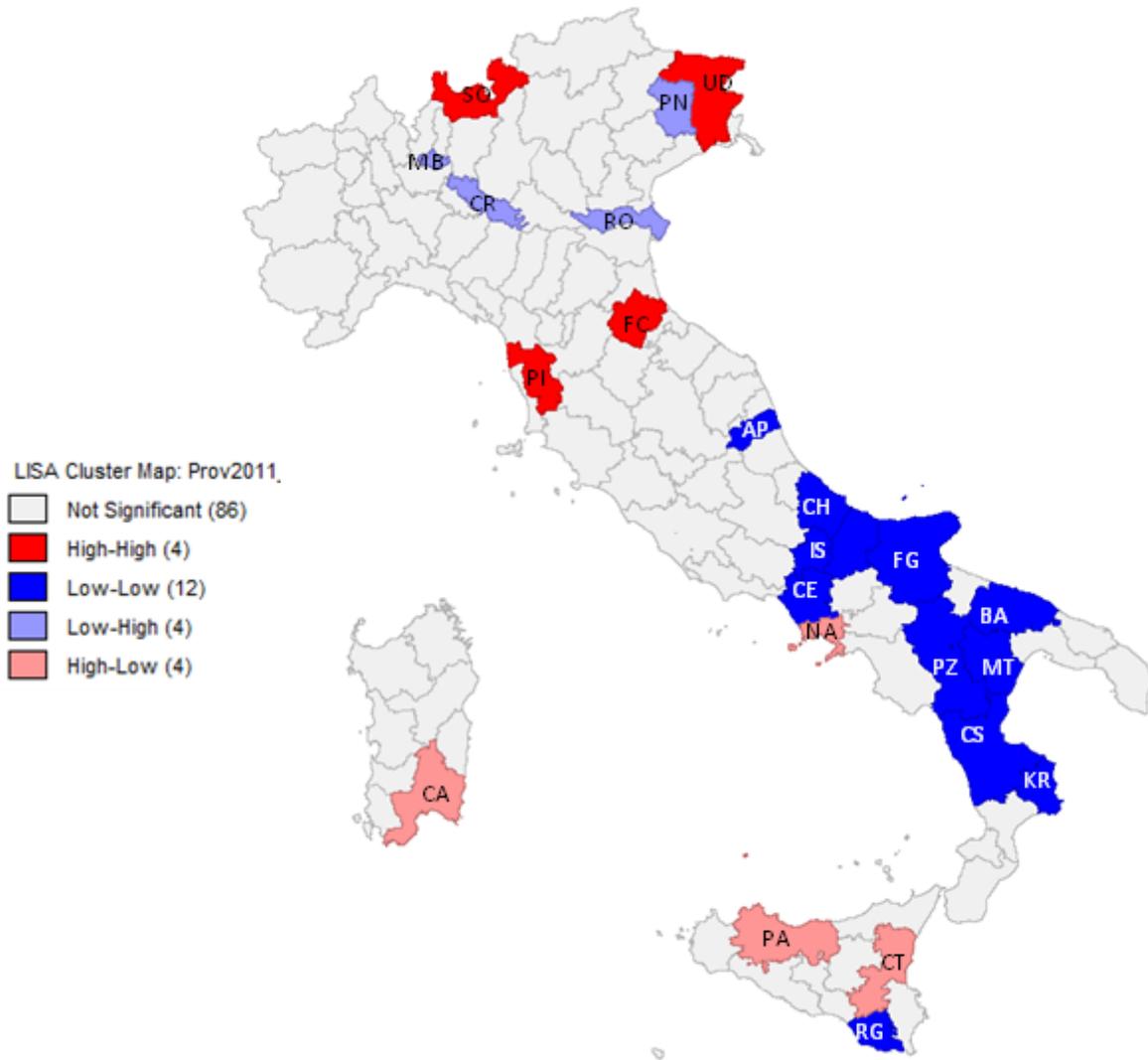
Figure 3

Moran's Index.

Figure 4

Spatial autocorrelation - LISA Cluster map Index. (in square brackets the number of provinces)