# Analyzing clinical processes and detecting potential correlation between CKD and air pollution[1]

*Erika Pasceri*
University of Calabria, Italy
*Anna Perri*
University of Calabria, Italy
*Giovanna Aracri*
Institute for Informatics and Telematics – CNR, Italy
*Sergio Cinnirella*
Institute of Atmospheric Pollution Research – CNR, Italy

## *Abstract*

Noncommunicable diseases (NCDs) are the most common causes of morbidity and premature mortality worldwide. The National Chronicity Plan of the Italian Ministry of Health classified Chronic Kidney Disease (CKD) as a chronic disease with a significant level of criticality, highlighting that its management has to be conducted through information systems. To this aim, starting from the clinical data collected within the database of the Department of Nephrology, Dialysis and Transplantation (Annunziata Hospital – Cosenza – Italy), we carried out a propaedeutic analysis of the clinical procedures which are commonly adopted in the care process of CKD defining all procedural flows through process modelling techniques. Furthermore, all clinical data analysed, as it has been reported that environmental pollutants can potentially increase the risk of CKD or accelerate its progression, have been used to define a distribution map over the Calabria Region in order to detect the potential correlation between CKD and air pollution.

## *1.0 Introduction*

Knowledge Organization Systems (KOSs) are conceptual structures, which collect standardized terminology that allow managing, retrieving and sharing data, information and knowledge. They traditionally have been built and maintained throughout the world to support subject indexing of li-

---

[1] Authors have equally contributed to this work, however Giovanna Aracri particularly focused on "Introduction", Erika Pasceri on "Method", "Discussion" and "Conclusion"; Sergio Cinnirella on "CKD data", "Pollution Data", "Data Integration"; Anna Perri on "Conclusion".

brary resources (such as books, journals, manuscripts, videos, etc.). However, they can also be used for many other functions (e.g. Information Retrieval (IR), domain knowledge representation and organization, metadata assignment, etc.) (Zeng 2008) and for designing sophisticated platforms able to detect, discover and integrate knowledge (Soergel 2009). The acronym KOSs coined by the Networked Knowledge Organization Systems Working Group (NKOS) gathers a set of controlled vocabularies (terminologies, classification systems, thesauri, semantic networks, ontologies, etc.). Their aim is to systematize and make explicit the semantic structure of both: the general knowledge intended as the range of all known subjects and disciplines; and the domain specific knowledge particularly referred to the set of knowledge acquired through studies or experience by experts in a community of practice. The KOSs optimize the effectiveness of the search systems because they actively contribute to increase the degree of Precision and Recall measures, in order to reduce noise due to non-relevant results and improve their accuracy and pertinence. Thus, the combination of KOSs with IR strategies guarantees the match of the specific request made by the user and the real set of information stored within a collection, a database, a platform, etc. As is known, the IR techniques in order to achieve their objectives need to convert the information request formulated in a natural language into a more formal, precise and unambiguous language. Therefore, this means that both the syntax and the semantics are codified and transmitted according to the latest standard in order to accomplish interoperability. The syntax is essential for implementation and control purposes because it ensures to structure data and information by using reliable, good quality and modeled metadata. On the contrary, the semantics deals with word sense and meaning and therefore with the content of metadata. Syntax and semantics are strictly linked. Their combination is very important since it reduces both lexical and syntactic ambiguity due to the use of synonyms or quasi-synonyms, polysemy and homonyms (ISO 25964-1:2011). Automatically detecting and extracting data and information according to specific parameters and attributes in some domain is crucial. Therefore, the combined use of flexible formats such as Extensible Markup Language (XML) and Unified Model Language (UML) and, KOS especially in some domains, where an extremely high quality of data and information is required, is important for generating reliable sources of data and information. KOSs are useful to support indexing, IR, Knowledge organization and representation and metadata assignment and to accomplish further scopes concerning data and information extraction and integration coming from heterogeneous sources and referred to different but closely related knowledge domains, as we envisage to highlight in this paper. The core of this study regards the definition of the procedural flows about Chronic Kidney Disease (CKD) and the analysis of data collected in order to explore potential correlation between CKD and air pollution.

The huge amount of clinical data produced and stored everyday by health providers needs to be collected and managed in a precise way. This enables data to be accurately used, firstly by the scientific community to constantly learn from each other, especially when the knowledge grows so quickly as in the clinical domain, and subsequently in clinical practice (Cardillo 2016). The relevance of using KOSs in the healthcare domain is widely discussed in the literature. This is strictly related to vocabularies, terminologies or classification and coding systems used to better organize and define clinical concepts and to identify access keys to codified data that can thus be combined, manipulated and shared among healthcare professionals (physicians, data analysts, and all the healthcare operators) involved in the process of care. The KOSs in this way allow to structure and represent complex information fostering their correct interpretation and sharing and to use them in an interoperable way (Cardillo et al. 2016). Nowadays information has been the key to a better organization and new developments. The more information we have, the better we can optimally organize data to deliver the best outcomes (Dash et al. 2019). That is particularly true in the clinical because data and in turn information is used to an improve prediction of a particular disease on the basis of certain parameters. The authors analyzed big data in healthcare, highlighting the management and future prospects, stating that:

> “Healthcare is a multi-dimensional system established with the sole aim for the prevention, diagnosis, and treatment of health-related issues or impairments in human beings. The major components of a healthcare system are the health professionals (physicians or nurses), health facilities (clinics, hospitals for delivering medicines and other diagnosis or treatment technologies), and a financing institution supporting the former two” (Dash et al. 2019, 3).

In each level of the entire care process there is always someone who enters and manages data according to the previous phase, so each professional is responsible for it. The outcome is strictly dependent on the accuracy and completeness of the entire process. The digitization of healthcare has strongly contributed to foster the development of these aspects. On the one hand the transposition of all processes (how data have been managed and transferred) into a digital environment that means to strictly apply rules according to standards adopted (such as HL7), on the other hand data integration that must be compliant to coding or classification systems in use (such as the International Classification of Diseases, Logical Observation Identifiers Names and Codes, Anatomical Therapeutic Chemical classification system, etc.). Concomitantly, the adoption of Electronic Health Record (EHR) systems to store and standardize medical and clinical data – slow at the beginning of 21st century but grown significantly after 2009 (Reisman 2017) – contribute to reduce significantly: the growth of additional or redundant examinations; ambiguities caused by handwriting; time and health waste for chronic patients and for the entire clinical management. Benefits in using EHRs on the clinician side lay on the chance to analyze the entire medical history of patients.

CKD represents a global public health issue, whose prevalence has gradually increased over the past decade because of emerging risk factors, including some environmental chemicals and particulate matter (PM) that worsen the renal function (Nugent et al. 2011; Tsai et al. 2021). The kidneys excrete waste products from the body, therefore they are susceptible to the adverse effects exerted by the toxins and pollutants circulating in the blood. Epidemiological evidences highlighted that environmental pollutants are important factors in the etiology of CKD and in vitro and in vivo studies elucidated some of the molecular mechanisms by which pollutants, in particular, the heavy metals and PM promote kidney damage (Tsai et al. 2021; Cardelis et al. 2014; Kim et al. 2015). The patients affected by CKD have a higher risk of progression to dialysis and cardiovascular mortality as well as cancer (Raaschou-Nielsen et al. 2017). In addition, long-term exposure to PM has been found responsible for damage to the glomerular barrier and the renal tubule, with consequent alteration of renal function (Lue et al. 2013; Xu et al. 2016). In fact, clinical studies have linked dust levels in the atmosphere with CKD (Bragg-Gresham et al. 2018). Hence, CKD is considered a one of the main health problems for the governments of all countries, whose costs are of great magnitude, as it consumes important percentages of the gross domestic product of the countries (Burgos-Calderón et al. 2021). Therefore, it is crucial to implement national prevention programs to reduce modifiable risk factors, such as exposure to environmental pollutants. Furthermore, the implementation of screening population programs is likewise essential for the early detection of kidney damage, as well as standardized protocols towards a better management of the patients affected by CKD, according to its stage. In this *scenario*, during the medical examination, the clinicians should perform a detailed exposure assessment to establish the potential nephrotoxicants exposure involved in the onset of the kidney disease and its progression to the CKD. More interestingly, the collected information should be encoded within the EHR and not be used only for diagnostic but also for epidemiological purposes.

## *2.0 The method*

The eHealth challenges contribute to the centrality of the inclusive approach to patient care. This awareness is rapidly gaining popularity because it generates a series of benefits concerning the quality of the treatment, the therapeutic and diagnosis choices, the reduction of the waiting time, the simplification of the patient management during his pathway care, especially when he is suffering from chronic pathology as CKD. As previously mentioned, the CKD causes a progressive and complete loss of the kidney function and its progression consists in five stages, each of which requires targeted interventions according to gravity and a multidisciplinary team for handling them in the best possible way. The involvement of several types of health professionals (nephrologists, psychologists and nutritionists) according to their specific competences and experience generates a diversified knowledge that needs to be collected, organized, managed and shared with all subjects taking part in the clinical pathway. From a strictly clinical viewpoint, the Diagnostic – Therapeutic – Assistance Paths (DTAP) listed in national and international guidelines and protocols, already provide a standardized framework for managing some certain patient categories. However, the current organization of the health services especially in the hospital context is rather fragmented. This condition can be due to both: the inadequate availability of clinical specialists and the difficulty of accessing and adequately using the tools already existing. In this sense, it is important to take into consideration, in addition to the clinical dimension, also the technical and the semantic ones. These latter aspects, despite their relevance, are too often neglected and considered as marginal in designing a clinical pathway. Indeed, mixing these layers in reengineering the workflow of a clinical process ensures quality of care throughout the course of the disease. This assumption is valid especially when the complexity of the disease is such as to require several expert's consultation and a close cooperation among them. Guidelines and protocols highlight that the performance improvement of a care process also depends on how information are represented, transmitted and shared. It is an important step towards a greater effectiveness and availability of a treatment. Using IT technologies and devices to support the clinical pathway makes a positive impact from different points of view. From a managerial and organizational perspective, benefits concerning the reduction of waiting times, the way in which information are clearly acquired and transferred, the availability and accessibility of a patient's clinical history; whereas from a purely clinical angle, their use supports the equity in treatment, disease monitoring, etc. In this section, we are going to describe the integrated workflow aimed at providing a clear and comprehensive vision of the CKD. A specific focus has been dedicated to the use of the International Classification of Diseases, 9th revision – Clinical Modification (ICD9-CM) codes adopted to identify diagnosis, services and procedures. The conceptualization of this management workflow allows the collection and the querying of large amounts of heterogeneous data in order to support the development of a modular and scalable platform with the ambitious goal of transforming data into knowledge. It is, therefore, essential to plan the clinical activity by using tools and strategies able to manage, monitor and report all the information produced, whenever a patient accesses a health facility in the form of outpatient service, day-hospital admission, and ordinary hospitalization. Readily available and updated information are synonymous with health services efficiency and quality and it is advantageous for doctors and patients alike. From the point of view of document management, the practices in use at the Department of Nephrology, Dialysis and Transplantation (Annunziata Hospital – Cosenza – Italy) were rather fragmented since they were hybrid. In the face of digital documents enriched by metadata and created by using suitable tools and applications several other documents were in paper format. Because of the lack of homogeneity, it was not possible to guarantee a clear information interpretation to determine statistics or check some regularities of the disease. The attempt was to reengineer the information workflow in use and make it interoperable and compliant to national and international standards by considering step-by-step the actors involved in the document type and format and the terminology and nomenclatures. The combination of these three variables allow establishing: who produces, consults, and edits a document; which document is subject to undergoing revisions and which value set has been modified. Documents produced or received by the Department of Nephrology and included in the information workflow may be medical prescriptions, tickets, clinical reports and discharge letters; each of these can be produced inside or outside the hospital by one or more medical specialists. Therefore, in order to support interoperability and make clinical documents and data searchable, it is necessary to employ international clinical standards such as Health Level 7 (HL7), ICD9-CM, Logical Observation Identifiers Names and Codes (LOINC) etc.

To accomplish the general goal, clinical data of 5017 patients coming from the Department of Nephrology's database have been analyzed. The focus was on patients, whose diagnosis has been codified by means of the 585 ICD9-CM code *Chronic kidney disease*. The following steps have been dedicated to build a database containing clinical and air pollution data with the aim to verify any type of correlation among them (Figure 1).

## 2.1 CKD data

As stated before CKD data have been collected from the database of Department of Nephrology. This basic information included for each patient the place and date of birth, place of residence and any lifestyle issues. All information has been collected strictly in anonymous form and authorized by

*Figure 1. Workflow adopted to integrate domain-specific datasets used to establish the correlation between particulate matter pollution and CKD*

the Ethics Committee of the Annunziata Hospital in order to comply with the current legislation on the protection of personal data.

The harmonization process consisted in the preparation of a dataset to proceed to the subsequent geocoding phase. In detail and where necessary, addresses were fixed such as follow for example:

*from* Via Milano n ° 1 *to* Via Milano, 1
*from* Via R. Salerno, 8 *to* Via Rosario Salerno, 8
*from* Via Kennedi *to* Via Kennedy

The geocoding process consisted in the transformation of the address name to the geographical position (latitude and longitude). To this aim the open source Nominatim application[2] that searches OpenStreetMap data by address and generates points coordinates was used.

## 2.2 Pollution data

A PM dataset was collected from NASA's data center Earth Observing System Data and Information System (https://beta.sedac.ciesin.columbia.edu). The historical PM outdoor trend (1998-2016) is reported as Global Annual $PM_{2.5}$ Grids retrieved from MODIS, MISR and SeaWiFS satellites. It consists of annual concentrations (micrograms per cubic meter) of ground-level fine PM (PM with dimension less than 2.5 micron) at a ground resolution of 0.01 degrees (van Donkelaar et al. 2016; van Donkelaar et al. 2018).

In addition, to consider indoor pollution, wood consumption by different climatic belts was calculated following the methodology proposed by Ozgen et al. (2014) and using the last available socio-economic dataset

---

[2] "Nominatim," Github, last accessed October 22, 2021, https://github.com/osm-search/Nominatim.

(http://dati-censimentopopolazione.istat.it/Index.aspx): resident population in built-up areas, residential dwellings, number of households, number of residential buildings and data on the consumption of fuel wood.

## 2.3 Data Integration

By means of QGIS (https://www.qgis.org) an open source Geographic Information System (GIS), specific values of PM at different locations were obtained. The integrated dataset was finally superimposed to a Digital Terrain Model (DTM) to obtain altitudes at each location. The final dataset was finally prepared as a database for future possible elaborations and an example is reported in Table 1:

| Variable | Description | Unit of Measure | Example |
|---|---|---|---|
| id_orig | Dataset Id | # | 1 |
| id_paz | Patient Id | # | 9877-41 |
| provenienza | City of origin | text | Acquaformosa |
| eta | Age | years | 59 |
| com_nascit | Born place | text | Lungro |
| com_resid | Residence | text | Acquaformosa |
| cap_resid | Postal code | | 87010 |
| lat_deg | Latitude | Decimal degrees | 39.722071 |
| lon_deg | Longitude | Decimal degrees | 16.090241 |
| cod_reg | National stistic code – Region | # | 18 |
| regione | Region | text | Calabria |
| cod_pro | National stistic code – Province | # | 78 |
| provincia | Province | text | Cosenza |
| cod_com | National stistic code – Municipality | # | 1 |
| comune | Municipality | text | Acquaformosa |
| cod_loc | National statistic code – fraction | # | 10001 |
| localita | Fraction | text | Acquaformosa |
| quota_m | Altitude | meters | 781 |
| pop_2000_# | Population in 2000 | # | 1454 |
| pop_2011_# | Population in 2011 | # | 1145 |

| Variable | Description | Unit of Measure | Example |
|---|---|---|---|
| abit_resid_# | Number of houses | # | 480 |
| fam_resid_# | Number of families | # | 480 |
| edifici_# | Total buildings | # | 693 |
| edifici_usati_# | Inhabited buildings | # | 682 |
| edifici_residenti_# | Residential buildings | # | 574 |
| firewood_cons_t | Firewood consumption | tons | 24.96 |
| energy_MJ | Energy produced by firewood | Mega Joule | 481728 |
| PM_indoor_ug | PM from firewood combustion | µg/m³ | 246.645 |
| pm25_1998_ug | Average PM, 1998 | µg/m³ | 6 |
| pm25_1999_ug | Average PM, 1999 | µg/m³ | 7.2 |
| pm25_2000_ug | Average PM, 2000 | µg/m³ | 5.6 |
| pm25_2001_ug | Average PM, 2001 | µg/m³ | 6.1 |
| pm25_2002_ug | Average PM, 2002 | µg/m³ | 6.1 |
| pm25_2003_ug | Average PM, 2003 | µg/m³ | 7.6 |
| pm25_2004_ug | Average PM, 2004 | µg/m³ | 5.2 |
| pm25_2005_ug | Average PM, 2005 | µg/m³ | 6 |
| pm25_2006_ug | Average PM, 2006 | µg/m³ | 5.7 |
| pm25_2007_ug | Average PM, 2007 | µg/m³ | 4.6 |
| pm25_2008_ug | Average PM, 2008 | µg/m³ | 4.1 |
| pm25_2009_ug | Average PM, 2009 | µg/m³ | 4.7 |
| pm25_2010_ug | Average PM, 2010 | µg/m³ | 4.1 |
| pm25_2011_ug | Average PM, 2011 | µg/m³ | 4.8 |
| pm25_2012_ug | Average PM, 2012 | µg/m³ | 4.9 |
| pm25_2013_ug | Average PM, 2013 | µg/m³ | 3.9 |
| pm25_2014_ug | Average PM, 2014 | µg/m³ | 4.1 |
| pm25_2015_ug | Average PM, 2015 | µg/m³ | 5.9 |
| pm25_2016_ug | Average PM, 2016 | µg/m³ | 4.9 |

*Table 1. List and example of variables included in the dataset*

## *3.0 Results*

The examined dataset consisted of 5017 patients (records) mainly distributed in the Cosenza Province. The analysis using the heatmap methodology

*Figure 2. Heatmap of patients' distribution in the Province of Cosenza*

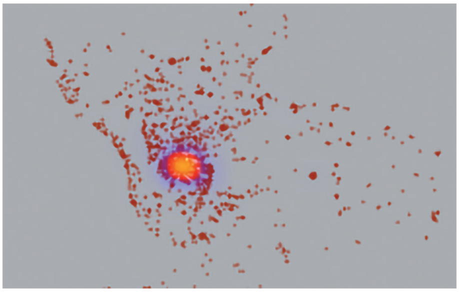

*Figure 3. Trend to cluster of CKD patients*

(Wilkinson and Friendly 2009) shows some density clusters in the Cosenza area (circular area with a radius of 40 km around the center of Cosenza). The heatmap is a visual representation of data aggregation that encodes the aggregation of information using colors. Figure 2 shows the heatmap of the Province in which some clusters of expected density are highlighted for the urban area and surrounding large towns as Acri and Bisignano.

The analysis of the clustering tendency according to the Hopkins methodology (Hopkins and Gordon 1954) shows a strongly concentrated data structure in the Cosenza area ($h_{mean} = 0.99$) (Figure 3). By carrying out the Optics Clustering (minPoints = 150 and eps = 2450) the data were aggregated to form a single cluster focused on the urban area (Figure 4). By reducing the minPoints value to 70 it is possible to identify minor clusters already highlighted by the heatmap of Figure 5.

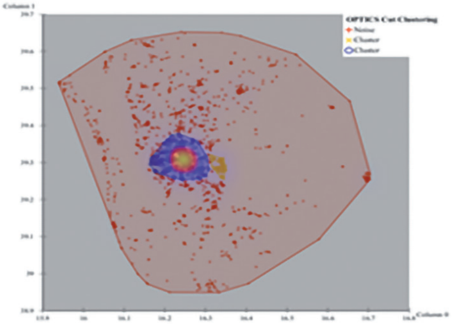

*Figure 4. Optics clustering of CKD patients. Clusters identified with minPoints = 150*

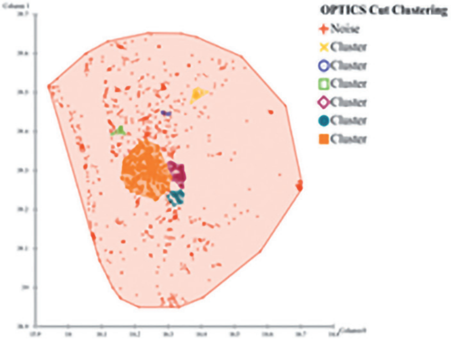

*Figure 5. Optics clustering of CKD patients. Clusters identified with minPoints = 70*

A further analysis concerned the distribution of particulate matter. Also in this case a heatmap was produced, which highlighted different levels of pollutants in the various areas (Figure 6). The figure shows as an example the distribution of the concentration for the year 2016.

## *4.0 Discussion*

The effective and complete application of EHRs is an essential prerequisite for the evolution of healthcare systems. It is also known that advancing the notion of clinical data as a public good and a central common resource for advancing knowledge is an evidence for effective care. In addition to this,

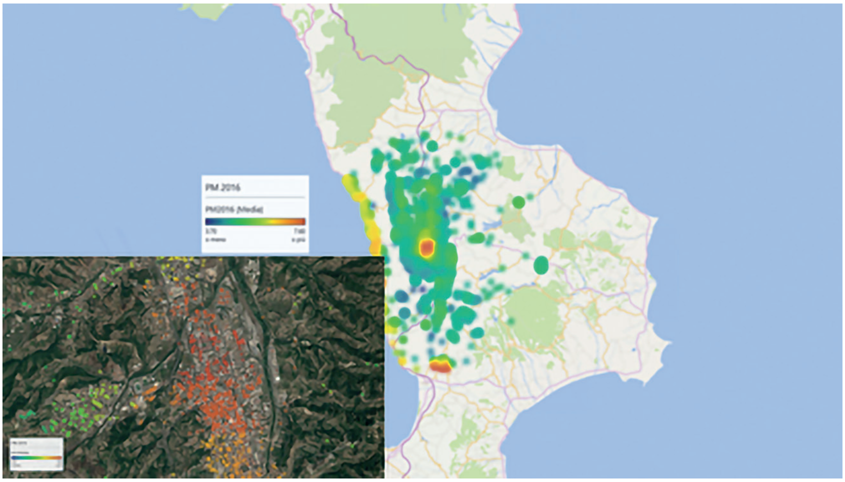

*Figure 6. Regional heatmap of the distribution of particulate matter and detail for Cosenza City*

obtaining database linkage, data useful for mining and interoperable patient record platforms to foster more rapid learning could better foster consistency and coordination in efforts to better manage knowledge (Institute of Medicine 2007).

In this way clinical decision support systems are considered to be an important vehicle for implementing new evidence and knowledge into daily practice. They generate patient-specific recommendations by matching individual patient characteristics to a knowledge base (Haynes 2010).

Encoded data on diagnoses and procedures are crucial to the success of the health system. Undoubtedly, the correct use of coding systems in the health field supports the management of health information for evidence-based decision-making, primarily for the chronic diseases. Moreover, it allows sharing and comparison of health information between different health facilities, ensuring an appropriate patient care process.

The analysis of the relationship between air pollution and CKD was carried out by considering the distribution of patients and the concentration of particulate matter in the Province of Cosenza. The results obtained by means of statistical analyses highlighted, as expected, the tendency to data aggregation in the urban area of Cosenza but, at the same time, the presence of unexpected clusters such those in surrounding towns. Even in the case of the distribution of particulate matter, clustering in the urban area is evident. The health and environmental data integration can provide unexpected challenges to identify environmental exposures to pollution that may affect human health (Bassil et al. 2015). Data integration and access can enable more punctual decisions through powerful knowledge platforms that bring scientific information prone to the general public and decision-makers.

In addition, the cause and severity of how different environmental pollution and exposures impact human health can be better understood if such integration will be increased, also by fostering data harmonization. There is no perfect methodology to link health and environmental data (IEHIAS, 2021) but the use of GIS is a first step despite its incomplete provision of knowledge (Tim 1995; Sipe and Dale 2003). Much effort is required especially in the collection of harmonized health information that often suffer from fragmentation, clarity and collection methodology (Geneviève et al. 2019). This harmonization and integration can improve data quality, re-usability and interoperability.

## *5.0 Conclusion*

In this study we have investigated the correlation between CKD and air pollution, focusing our attention on $PM_{2.5}$, as experimental and clinical studies demonstrated that the exposure to air pollution, especially PM with a diameter smaller than 10 µm, impairs renal function and increases the risk of incident CKD, leading to end-stage-renal-diseases (Feng et al. 2021). However, several studies demonstrated that other environmental pollutants, including metals and phthalate, exert nephrotoxic effects, potentially increase the risk of CKD or accelerate its progression (Tsai et al. 2021; Wu 2020). In addition, some studies link CKD to intake of metals such as arsenic, cadmium, lead, mercury, copper which have a high nephrotoxic impact (Orr and Bridges 2017; Sun et al. 2019).

The analyses on both environmental pollution and CKD distribution, despite matching, cannot lead to the conclusion that there is a direct relationship. The most significant missing information is the temporal exposure of patients to PM. Moreover, information on metals and phthalate can better address the analysis on CKD-air pollution causal links. Such information not available in our case does not allow a final conclusion on the impact of PM on CKD. Therefore, further studies will be conducted on a larger sample of patients from different areas of the Calabria region, to investigate the correlation of CKD with the environmental pollutants locally detected.

Moreover, an extensive and precise use of coding could better support the decision support systems in analysis focused on heterogeneous data. On the one hand, by using more ICD9-CM specific codes (i.e., 585.3 Chronic kidney disease, Stage III (moderate) to detect specific classes of patients; on the other hand, by using other clinical codings in a complementary manner, for example the LOINC for laboratory reports and the ATC classification system, to detect more information about drug treatments, allergy etc.

## *References*

Bassil, Kate L., Margaret Sanborn, Russ Lopez and Peter M. Orris P. 2015. "Integrating Environmental and Human Health Databases in the Great Lakes Basin: Themes, Challenges and Future Directions." *International Journal of Environmental Research and Public Health* 12, no. 4: 3600–14. https://doi.org/10.3390/ijerph120403600.

Bragg-Gresham, Jennifer, Hal Morgenstern, William McClellan, Sharon Saydah, Meda Pavkov, Desmond Williams, Neil Powe, Delphine Tuot, Raymond Hsu, Rajiv Saran, et al. 2018. "County-level air quality and the prevalence of diagnosed chronic kidney disease in the US Medicare population." *PloS One* 13, no. 7 e0200612. 31. https://doi.org/10.1371/journal.pone.0200612.

Burgos-Calderón, Rafael, Santos Á. Depine, and Gustavo Aroca-Martínez. 2021. "Population Kidney Health. A New Paradigm for Chronic Kidney Disease Management." *International Journal of Environmental Research and Public Health* 8, no. 13: 6786. https://doi.org/10.3390/ijerph18136786.

Cardillo, Elena, Maria Teresa Chiaravalloti and Erika Pasceri. 2016. "Healthcare Terminology Management and Integration in Italy: Where we are and What we need for Semantic Interoperability." *European Journal for Biomedical Informatics* 22, no. 1.

Dash, Sabyasachi, Sushil Shakyawar, Mohit Sharma, et al. 2019. "Big data in healthcare: management, analysis and future prospects." *J Big Data* 6, no. 54. https://doi.org/10.1186/s40537-019-0217-0.

Feng, Ying-Mei, Lutgarde Thijs, Zhen-Yu Zhang, Esmée M. Bujnens, Wen-Yi Yang, Fang-Fei Wei, Bram G. Janssen, Tim S. Nawrot, and Jan A. Stanssen. 2021. "Glomerular function in relation to fine airborne particulate matter in a representative population sample." *Scientific Reports* 11, no. 14646. https://doi.org/10.1038/s41598-021-94136-1.

Geneviève, Lester D., Andrea Martani, Maria C., Mallet Tenzin Wangmo, and Bernice S. Elger. 2019. "Factors influencing harmonized health data collection, sharing and linkage in Denmark and Switzerland: A systematic review". *PLoS One* 14, no. 12: e0226015. https://doi.org/10.1371/journal.pone.0226015.

Haynes, R. Brian, and Nancy L. Wilczynski. 2010. "Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: Methods of a decision-maker-researcher partnership systematic review." *Implementation science: IS* 5, no. 12. https://doi.org/10.1186/1748-5908-5-12.

Hopkins, Brian, and Gordon John Skellam. 1954. "A new method for determining the type of distribution of plant individuals." *Annals of Botany* 18, no. 2: 213–27.

IEHIAS, 2021. "Integrated Environmental Health Impact Assessment System." Last accessed October 22, 2021. http://www.integrated-assessment.eu.

Institute of Medicine. 2007. *The Learning Healthcare System: Workshop Summary*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11903.

ISO 25964-1:2011, *Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval*.

Kim, Ki-Hun, Kabir Ehsanul, and Shamir Kabir. 2015. "A review on the human health impact of airborne particulate matter." *Environment international* 74:136-43.

Maddugari, Santosh Kumar, Vijaay B. Borghate, Sidarth Sabyasachi, and Raghavendra Reddy Karasani. 2019. "A Linear-Generator-Based Wave Power Plant Model Using Reliable Multilevel Inverter." In *IEEE Transactions on Industry Applications*, vol. 55, no. 3: 2964-72. https://doi.org/10.1109/TIA.2019.2900604.

Mehta, Amar J., Antonella Zanobetti, Marie-Abele C. Bind, Itai Kloog, Petros Koutrakis, David Sparrow, Pantel S. Vokonas, and Joel D. Schwartz. 2016. "Long-term exposure to ambient fine particulate matter and renal function in older men: The Veterans Administration Normative Aging Study." *Environmental Health Perspectives* 124, no. 9: 1353-60.

"Nominatim." Github, Last accessed October 22, 2021. https://github.com/osm-search/Nominatim.

Nugent, Rachel, Fathima Sana F., Andrea Feigl, and Dorothy Chyung. 2011. "The burden of chronic kidney disease on developing nations: a 21st century challenge in global health." *Nephron Clinical practice* 118, no. 3: c269-77.

Orr, Sarah E., and Christy C. Bridges. 2017. "Chronic Kidney Disease and Exposure to Nephrotoxic Metals" *International Journal of Molecular Sciences* 18, no. 5: 1039. https://doi.org/10.3390/ijms18051039.

Ozgen, Senem, Stefano Caserini, Silvia Galante, Michele Giugliano, Elisabetta Angelino, Alessandro Marongiu, Francesca Hugony, Gabriele Migliavacca, and Morreale Carmen. 2014. "Emission factors from small scale appliances burning wood and pellets." *Atmospheric Environment* 94: 144-53. https://doi.org/10.1016/j.atmosenv.2014.05.032.

Reisman, Miriam. 2017. "EHRs: the challenge of making electronic data usable and interoperable." *P & T: a peer-reviewed journal for formulary management* 42, no. 9: 572-75.

Raaschou-Nielsen, Ole, Marie Pedersen, Massimo Stafoggia, Gudrun Weinmayr, Zorana J. Andersen, Claudia Galassi, Johan Sommar, Bertil Forsberg, David Olsson, Bente Oftedal, et al. 2017. "Outdoor air pollution and risk for kidney parenchyma cancer in 14 European cohorts". *International journal of cancer* 140, no. 7: 1528-1537. https://doi.org/10.1002/ijc.30587.

Sipe, Neil and Pat Dale. 2003. "Challenges in using geographic information systems (GIS) to understand and control malaria in Indonesia." *Malaria Journal* 2. https://doi.org/10.1186/1475-2875-2-36.

Soergel, Dagobert. 2009. "Knowledge Organization Systems. Overview".

Sun, Yaofei, Quan Zhou, and Jie Zheng. 2019. "Nephrotoxic metals of cadmium, lead, mercury andrsenic and the odds of kidney stones in adults: An exposure-response analysis of NHANES 2007–2016." *Environment International* 132:105115. https://doi.org/10.1016/j.envint.2019.105115.

Tim U.S. 1995. "The Application of GIS in Environmental Health Sciences: Opportunities and Limitations." *Environmental Research*, 71, no. 2: 75-88.

Trigari, Marisa. 1993. *Come costruire un thesaurus*. Franco Cosimo Panini.

Tsai, Hui-Ju, Pei-Yu Wu, Jiun-Chi Huang, and Szu-Chia Chen. 2021. "Environmental Pollution and Chronic Kidney Disease." *International Journal of Medical Science* 18, no. 5:1121-29. https://doi.org/10.7150/ijms.51594.

van Donkelaar, Aaron, and Randall V. Martin. 2016. "Global Estimates of Fine Particulate Matter Using a Combined Geophysical-Statistical Method with Information from Satellites." *Environmental Science & Technology* 50, no. 7: 3762-72. https://doi.org/10.1021/acs.est.5b05833.

van Donkelaar, Aaron, Randall V. Martin, Micheal Brauer, N. Christina Hsu, Ralph A. Kahn, Robert C. Levy, Alexei Lyapustin, Andrew M. Sayer, and David M. Winker. 2018. "Global Annual PM2.5 Grids from MODIS, MISR and SeaWiFS Aerosol Optical Depth (AOD) with GWR, 1998-2016." Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4ZK5DQS.

Wilkinson, Leland, and Micheal Friendly. 2009. "The History of the Cluster Heat Map." *The American Statistician* 63, no. 2: 179-84.

Xu, Xin, Guobao Wang, Nan Chen, Tao Lu, Sheng Nie, Gang Xu, Ping Zhang, Yang Luo, Yongping Wang, Xiaobin Wang, et al. 2016. "Long-term exposure to air pollution and increased risk of membranous nephropathy in China." *Journal of the American Society of Nephrology*: *JASN* 27, no. 12: 3739-46. https://doi.org/10.1681/ASN.2016010093.

Zeng, Marcia Lei. 2008. "Knowledge Organization Systems (KOS)." *Knowledge Organization*, 35, nos 2/3:160 -182. https://doi.org/10.5771/0943-7444-2008-2-3-160.

Wu, Mei-Yi, Wei-Cheng Lo, Chia-Ter Chao, Mai-Szu Wu, and Chih-Kang Chiang. 2020. "Association between air pollutants and development of chronic kidney disease: A systematic review and meta-analysis". *Science of the Total Environment* 706: 35522.