DISEASE MAPPING AND RISK ASSESSMENT FOR PUBLIC HEALTH DECISION-MAKING

Report on a WHO Workshop

Rome, Italy
2–4 October 1997
ABSTRACT

Spatially referenced epidemiological data and analytical methods recently developed for geographical studies are becoming increasingly common. Although in-depth investigations can be carried out, problems remain with interpretation and with evaluation of the implications. The Workshop aimed at clarifying the relevance of geographical studies and their role in informing public health decision-making. Several issues of interest were discussed, including the ability of geographical studies to assess the causality of observed associations; to estimate their magnitude, uncertainty and dose–response relationship; to evaluate whether effects apply to individuals or populations; to communicate the findings from geographical studies in epidemiology between specialists, decision-makers, the media and the public; and to assess the need and effectiveness of public health action or further investigation following sporadic disease outbreaks. The Workshop also covered analytical methods for disease mapping, cluster investigation, ecological analyses, studies of risk near point sources of environmental pollution, and surveillance. The available methodology was reviewed and some original methods were presented and discussed in special sessions, where examples of applications were given.

Keywords

PUBLIC HEALTH
RISK ASSESSMENT
DECISION-MAKING
GEOGRAPHY
EPIDEMIOLOGICAL METHODS
EVALUATION STUDIES
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>OBJECTIVES</td>
<td>1</td>
</tr>
<tr>
<td>SUMMARY OF FIVE PARALLEL SESSIONS (CONTRIBUTED PAPERS)</td>
<td>2</td>
</tr>
<tr>
<td><strong>SESSION 1: DISEASE MAPPING</strong></td>
<td>2</td>
</tr>
<tr>
<td><em>Chairperson: Noel Cressie</em></td>
<td></td>
</tr>
<tr>
<td><strong>SESSION 2: CLUSTERING</strong></td>
<td>3</td>
</tr>
<tr>
<td><em>Chairperson: Julian Besag</em></td>
<td></td>
</tr>
<tr>
<td><strong>SESSION 3: ECOLOGICAL ANALYSIS</strong></td>
<td>4</td>
</tr>
<tr>
<td><em>Chairperson: Arnoldo Frigessi</em></td>
<td></td>
</tr>
<tr>
<td><strong>SESSION 4: RISK ASSESSMENT AROUND PUTATIVE SOURCES</strong></td>
<td>5</td>
</tr>
<tr>
<td><em>Chairperson: Göran Pershagen</em></td>
<td></td>
</tr>
<tr>
<td><strong>SESSION 5: HEALTH SURVEILLANCE AND APPLICATIONS</strong></td>
<td>7</td>
</tr>
<tr>
<td><em>Chairperson: Benedetto Terracini</em></td>
<td></td>
</tr>
<tr>
<td>CONCLUSIONS AND RECOMMENDATIONS</td>
<td>8</td>
</tr>
<tr>
<td>ANNEX 1. PARTICIPANTS</td>
<td>11</td>
</tr>
<tr>
<td>ANNEX 2. WORKING PAPERS</td>
<td>15</td>
</tr>
</tbody>
</table>
INTRODUCTION

In recent years many new methods have been developed in the field of disease mapping, and a number of health-oriented institutions from many countries have undertaken the production of atlases of diseases and mortality. Despite advances in methodology and increasing data availability, no systematic evaluation of the available techniques, with regard to their use in public health and decision-making, has been done so far. Besides disease mapping and descriptive studies, a growing interest has emerged on the evaluation of putative sources of risk. Initiated by the querelle on the risk of childhood leukaemia around nuclear plants to the recent claims on raised risk of asthma in areas with high traffic pollution, a variety of methods for the analysis of case event clustering and their relation to sources of noxious agents have appeared in both statistical and epidemiological literature. Furthermore, some countries host research centres devoted to the investigation of claims regarding spontaneous clusters of disease. Thus, it was felt that it was urgently needed to clarify the conditions of use and the merit of different techniques to address such questions and to inform public health response in an appropriate way.

OBJECTIVES

The European Initiative in Disease Mapping and Risk Assessment and the WHO European Centre for Environment and Health, Rome Division organized an international workshop with the following aims:

- to review and assess the current development of methods of data analysis to be used in geographical epidemiological studies;
- to provide an evaluation of the application of each of the available approaches for public health use;
- to reach a consensus upon a list of recommendations on the use of the techniques that are most appropriate to orient public health policy decisions.

Specific areas of interest include: disease mapping and its role in health surveillance and public health resource allocation; ecological analyses and their controversial use in etiological research; the role of cluster detection in epidemiology; and the analysis of risk around putative sources, with special emphasis on environmental causes of diseases.

Some 37 temporary advisers, 9 observers and 4 WHO officers from 13 countries were invited to attend the workshop (see Annex 1 for a complete list of names and addresses). Professor Benedetto Terracini was elected Chairperson, Professor Noel Cressie Vice-Chairperson and Dr Marco Martuzzi Rapporteur.

---

1 In 1996, a European Initiative in Disease Mapping and Risk Assessment (DMRA, project coordinator A.B. Lawson, University of Abertay Dundee) was funded by the European Union under the second Biomed Programme, section on Risk Factors of Occupational and Environmental Diseases. The participating countries are Belgium (E. Lesaffre, University of Leuven), France (J.F. Viel, University of Besançon), Germany (D. Böhning, Free University, Berlin), Italy (A. Biggeri, University of Florence) and the United Kingdom (A.B. Lawson, University of Abertay Dundee). The objectives of the DMRA are to provide a review of current disease mapping methods in member countries; to review the available methods for assessment of geographical variations in disease; and to assess, via Europe-wide applications, the most appropriate spatial methods.
A total of 36 working papers were presented and discussed in five sessions as under (Annex 2):

1. Disease Mapping, chaired by Noel Cressie (United States)
2. Clustering, chaired by Julian Besag (United States)
3. Ecological Analyses, chaired by Arnoldo Frigessi (Italy/Norway)
4. Risk Assessment around Putative Sources, chaired by Göran Pershagen (Sweden)
5. Health Surveillance and Applications, chaired by Benedetto Terracini (Italy).

The topics of discussion and conclusions of these sessions are summarized below.

Following the five sessions, three parallel working groups chaired by Annibale Biggeri (Italy), Tony Fletcher (United Kingdom) and Jean-François Viel (France), addressed a series of questions regarding geographical analyses and their applications in public health. Each working group drafted a list of recommendations for disease mapping and risk assessment for public health decision-making. Finally, in a meeting held in plenary, the conclusions reached separately by the three working groups were discussed, a consensus was sought and the workshop’s conclusions and recommendations were finalized. These are reported below, in the form of answers to the questions discussed by the working groups. A draft of the present report was circulated for comments after the meeting with all participants before publication.

The working papers presented during the workshop have been peer-reviewed both by participants in the workshop and by external consultants and will be collected and published in a book jointly edited by the DMRA initiative and the WHO European Centre for Environment and Health.

**SUMMARY OF FIVE PARALLEL SESSIONS (CONTRIBUTED PAPERS)**

**SESSION 1: DISEASE MAPPING**

*Chairperson: Noel Cressie*

Eight papers were presented: Lawson et al. (LEA); Cressie and Stern (CS); Besag and Knorr-Held (BK); Böhning and Schlattman (BS); Louis (L); Mollié (MO); Martuzzi and Hills (MH); and Lawson, Dreassi and Biggeri (LDB).

Disease mapping has two common uses: smoothing away noise to draw maps and assessing specific hypotheses concerning incidence. The presentations and discussion were almost exclusively concerned with the merits of various statistical models and analyses for disease incidence rates of small areas. However, one paper (LEA) did discuss the situation where case-event data (i.e. point patterns of case locations) are available.

In discussion, there was considerable enthusiasm for the development of case-event models and associated data analysis. However, participants recognized that in practice exact locations are often ambiguous (e.g. home versus workplace). An attempt might be made to build a point-level model for this ambiguity, or the point-level case-event data might be aggregated into small-area counts, with the possibility of ecological bias caused by the aggregation.
For the most part, the session was devoted to modelling and analysis of small-area count data. It was thought that the standard Poisson assumption on count data is appropriate. Some of the papers mentioned hypothesis tests on the set of area-specific rates or risk, and all papers put some form of mixing distribution on them. For example, BS assumed that the risks are independently and identically distributed according to a discrete distribution; their interest was in classification of the small areas. MH put a continuous gamma distribution for the risks; their interest was in the posterior distribution of its variance and how different it was from zero (the case of constant risk). In all other papers BK, L, MO and CS used multivariate log-normal distributions. One question of interest (largely unresolved) was whether these models are consistent at different levels of aggregation.

Spatial hierarchical models are useful for disease mapping. Participants did not agree about the choice of model at the second level of the hierarchy, but there was agreement that such mixing distributions are needed. Two dominant approaches were presented to statistical analysis, both developed in the framework of Bayesian statistics. The empirical Bayes method is where an attempt is made to estimate parameters of “prior” distributions using observed marginal distributions. The second approach is the fully Bayesian approach, where the prior and “posterior” distributions are obtained via Markov Chain Monte Carlo (MCMC) computations (after assessing appropriate convergence). MO mapped smoothed rates using posterior means; BK has a dynamic temporal component and mapped posterior medians; L estimated the edf and associated ranks using squared error loss; CS estimated extrema with special loss functions and developed Bayesian diagnostics; LDB considered data outside the region of interest as missing and mapped the resulting posterior means.

In conclusion, the Bayesian approach (empirical or full) is appealing because almost any question can be asked and addressed. Incorporating spatial dependence requires an extra one or two parameters and is a worthwhile insurance against mis-specification of regression effects. If it is found that the spatial component is not needed, a reanalysis might be done without it to obtain a more parsimonious model. Finally, it should be realized that if a request is made for smoothing and mapping, there is a good chance that the smoothed estimates will be used later (perhaps inappropriately) to assess specific hypotheses. The paper by L makes this problem clear and gives an example of this important issue.

SESSION 2: CLUSTERING

Chairperson: Julian Besag

Six talks were given in this session by (in order) Drs Jacquez, Kulldorff, Lawson, Schmidtman, Tango and Zöllner. With the exception of a paper on MCMC methods, based on Bayesian statistics, all the approaches adopted a frequentist standpoint.

The resolution of the data which the authors had in mind ranged from county level, where data are reasonably reliable but correspondingly the analysis is rather coarse, to census enumeration district (ED) and even case level, for both of which there is a greater danger of database problems caused, for example, by incompatibility between numerators (the cases) and denominators (those at risk). Potential incompatibilities include the fact that cases occur over a substantial period of time but that the at-risk population may be measured on a single census day and hence be susceptible to substantial migration effects (e.g. new housing projects); and that location may be ambiguous (e.g. it can defined as place of residence, birth, work or school).
In the talks, the purpose of a geographical analysis was to provide a p-value for evidence of overall clustering in the data, based on one or more test statistics assessed with respect to a particular reference distribution; or to identify apparent clusters presumably often with the intention of follow-up analysis by case-control or other methods if an identifiable putative cause was subsequently suggested; or both of these.

Various different approaches to such investigations were suggested and some authors included limited comparisons but there was no consensus on what might be a “best” method. This is not unreasonable; a method of analysis should be appropriate to the eventual purpose, to knowledge about the disease etiology, to the form of the available data, and so on. Calculation of p-values depended on analytical approximations or, more commonly, Monte Carlo simulation. There was some support for more extensive power studies against meaningful alternatives.

As regards p-values, several speakers and discussants noted that for some tests, both ancient and modern, the calculations recommended in the literature are generally invalid. This is a quite common fault for tests of clustering in cancer atlases. For example, in Moran’s test and similar ones, the traditional Gaussian approximation or simple randomization analysis, applied to incidence rates or ranks is typically incorrect: even if risk is constant, there should be a geographical pattern of incidence rates if high (e.g. urban) and low (e.g. rural) populations show a geographical pattern, as they usually do. The reason is that high and low rates predominate in the low population zones and this induces the pattern. Speakers noted that the incorrect methodology can be replaced by Monte Carlo versions based, for example, on multivariate hypergeometric distributions, although these become highly computationally intensive when large populations are involved. Sometimes the incorrect results are adequate in practice but this needs careful monitoring. The existence of the problem requires more publicity but was not the focus of the session.

Practical examples were given by most speakers and the Chairperson.

SESSION 3: ECOLOGICAL ANALYSIS

Chairperson: Arnoldo Frigessi

Papers were presented by Drs Bernardinelli, Best, Divino, Braga, Ferrandiz, Frigessi, Kelly and Langford and were mainly concerned with hierarchical Bayesian modelling and modelling spatial interaction in the presence of informative covariates measured with or without error.

Several case studies have been described in order to introduce and illustrate old and new methods, including, among others, avoidable mortality for asthma; the association between malaria and diabetes; mapping of prostate and lung cancer mortality rates; and ultraviolet light as a possible cause of skin cancer.

The talks and the discussion focused on the following issues:

1. Bayesian complex models allow a realistic and structured description of nature, in particular through incorporation of cause/effects features and measurement errors in covariates. However, statistical tools for model validation and criticism do not yet seem fully adequate, and sensitivity studies are important. Cooperation between statisticians and other fields of expertise is needed for building reliable models.
2. Models may seem sometimes too large and overparametrized; covariates may act as confounders; and there is a risk of overadjustment to watch out for.

3. The introduction of spatial smoothing and interaction among estimated spatial parameters are now standard in the presence of covariates. There seem to be three main rationales behind the introduction of terms describing spatial features into models.

(a) When it is believed that measured covariates and the selected model are appropriate, the spatial part is introduced in order to check a posteriori that it is not significant. If this is the case, the selected model explains the phenomena under study well enough.

(b) The spatial part is significant and reflects what is not explained by the covariates included in the regression model. It may capture spatial features related to covariates that cannot be measured (e.g. diet) or even defined. The presence of such spatial residual leaves something unexplained from the epidemiological viewpoint. However, keeping a spatial part in the model seems generally cautious.

(c) Presence of strong spatially structured effects. The spatial part is a fundamental feature of the data under study. For example, in the case of infectious diseases, vicinity is a truly explanatory factor. Such informative spatial models can also be introduced in order to investigate certain hypotheses, for example the presence of an infective agent in the etiology of childhood leukaemia.

4. Two new modelling inferential ideas were introduced:

(a) a new way to handle data collected on different scales/grids, avoiding the aggregation of information to the least detailed scale; and

(b) a new way of doing non-parametric inference in spatial auto models with covariates, when covariates modulate spatial interaction; non parametric methods are computationally feasible and results can be usefully compared to those obtained from parametric models.

5. There should be concern over the possible misinterpretation of disease maps. Maps without clear information on the underlying assumptions, models and approaches should not be delivered to public health authorities without careful consideration.

6. Disease maps are often interpreted “macroscopically”, i.e. looking for overall spatial trends or large areas at increased risk, with no interest in the resolution at which detailed information is conveyed, such as the exact boundaries among areas with different level of risk. When this is likely to be the case, models should try to incorporate higher level features like templates, in the spirit of Bayesian image analysis.

SESSION 4. RISK ASSESSMENT AROUND PUTATIVE SOURCES
Chairperson: Göran Pershagen

Papers were presented by Drs Biggeri, Bithell, Armstrong (in lieu of Dr Dolk), Maul), Pershagen and Waller.

One paper presented several methods for case-control analysis of risk around putative sources. An example was provided in which the area under study was divided into circular annuli with the postulated sources of pollution in the centre. This suggested that lung cancer risk was related to distance of residence to city centre, as well as to an incinerator in a multivariate analysis including individual information on smoking, occupation and air particulate levels.
Another paper described some exploratory methods for testing the uniformity of a risk surface over a specified region using a relative risk function. This implies that the relative risk at any given point represents the risk relative to a population-weighted average for the region as a whole. Using as an example childhood cancer data from England, the paper discussed how far to go in analysis and presentation of results from such exploratory activities in situations where there is no significant heterogeneity in the data.

One presentation dealt with the power of focused score tests under mis-specified cluster models. Focused tests assess whether cases are clustered around some pre-specified potential sources of hazard in the study area. Mis-specified score tests imply that tests defined for one type of cluster model are applied to clustering generated through a different model. One problem, which was discussed, is that proximity is often not a good proxy for true exposure and the importance of accurate environmental exposure measurements was stressed.

Sequential monitoring of low event rates was discussed in one paper. The intention is to consider time intervals in the appearance of cases, making it possible to perform a stepwise assessment of the putative source while monitoring the data in a prospective design. One problem is to adjust the critical threshold so as to maintain type I error rate under a pre-specified level of significance during the entire study. Another concern arises when the phenomenon under study is rare since this may lead to a less reliable model fitting and invalid asymptotic distribution of standard test statistics. The methodology was applied to leukaemia incidence near a French nuclear reprocessing plant and did not indicate spatial-temporal clustering within the space-time window under investigation.

A multinational European study on congenital anomalies near hazardous waste landfill sites was also described. This involved 21 sites in 15 areas and a total of more than 1000 cases and 2000 controls. Overall there appeared to be a 30% increase in risk associated with residence within three kilometres of the sites following adjustment for socioeconomic status, maternal age and year of birth. Various types of bias were discussed such as confounding by socioeconomic status, other industrial sites, differences in hospital ascertainment and migration but it was considered unlikely that this explained the results. The problems in interpretation were regarded as not principally statistical but related to the lack of evidence on exposure and plausible etiologic pathways.

Finally, one paper focused on the methodology for assessing lung cancer risks near point emission sources. The occurrence of lung cancer shows substantial geographic variation and ecological analyses have provided useful hints for explaining these findings, but are generally not sufficient for assessing causal relationships. For example, detailed evaluation of the role of ambient air pollution requires data on important risk factors for lung cancer, particularly regarding smoking, to adequately assess confounding and interactions. Poor characterization of exposure is a major problem. Ideally measurements of air pollutants should cover time periods relevant for disease etiology and the role of individual differences in activity patterns and mobility for exposure should be assessed. International collaboration is desirable in studies of lung cancer near point emission sources, in view of the often small populations with excessive exposure near each site and the possibility of combining data from several sites.
SESSION 5: HEALTH SURVEILLANCE AND APPLICATIONS  
Chairperson: Benedetto Terracini

Papers were presented by Drs Axelson, Becker, Fletcher, Neutra, Terracini, Viel and Walter.

Three presentations described geographical analyses of the occurrence of a number of cancer types in Germany and in Ontario and of pleural cancer in Italy. In Ontario, the availability of databases on the distribution of social, economic, behavioural, nutritional and other health indicators allowed known risk factors to be considered in the analyses. This greatly reduced the regional correlation of the residuals, thus increasing the specificity of the identification of areas requiring further special investigation.

The other four presentations related to the interaction of statisticians/epidemiologists investigating geographical patterns of disease and the rest of the society. Members of the group felt that great caution should be exercised in using risk estimates obtained from aggregated data in order to estimate risk at the individual level. Possible exceptions are unusual circumstances of highly specific associations with high attributable risks (e.g. asbestos and mesothelioma).

The following points were emphasized:

1. Maps can be understood by lay people provided they include any caveat or other consideration for their interpretation (including absolute numbers). Although scientific uncertainties can be shared with the lay people, scientists should not miss any opportunity to remind that lack of evidence for an effect does not correspond to evidence of lack of effect.

2. Studies aiming at assessing the effectiveness of different forms of communication of findings to the lay people are needed.

3. The pros and cons of several categories of geographical analyses deserve attention:
   (a) mapping of crude rates, standardized incidence rations (SIR) or even case numbers;
   (b) mapping of smoothed and model-adjusted rates and SIRs;
   (c) procedures for detecting or locating clusters of disease, either at the individual level or at the level of counties or greater;
   (d) ecological studies, i.e. studies of rates in a series of populations as a function of the prevalence of risk factors in those populations, either without accounting for spatial autocorrelation or accounting for it by several methods;
   (e) case-control or cohort studies using procedures accounting for the effects of any spatial autocorrelation;
   (f) procedures using computer mapping to calculate:
      • the location and extent of exposure to pollution (e.g. numbers of schools near agricultural fields);
      • the number of persons at risk of some exposure;
      • exposures to agents like traffic fumes or electromagnetic fields;
   (g) procedures using computer mapping to recognize a pattern of disease location suggesting a particular mode of disease transmission (e.g. linearity of disease excess suggests a polluted water grid).
The workshop provided detailed discussion of the technical issues and applications of the first five categories above (points (a) through (e)). However, in practical applications the choice of methods for addressing points (f) and (g) is also important.

1. Geographical analyses of pre-clinical endpoints - when feasible - have a potential for informing public health action.

2. Whenever geographical data on more than one risk factor are available, the combined effects should be investigated in terms of effect modification as well as confounding.

3. When prioritising interventions, mere numerical comparisons between risks are incorrect, in that they ignore other values, such as the difference between imposed vs. self-inflicted risks.

4. Several episodes worldwide indicate that the media may encounter difficulties in communicating to the public, in a balanced way, epidemiological findings and their interpretation in terms of reliability and risk estimates. The collection of case histories in this respect may help systematize and perhaps prevent misunderstandings between scientists and the lay press.

5. There is a need for a better knowledge of the societal context (including politicians, activists, unions, etc.) in which geographical studies are carried out and for analysing and publishing episodes in which the epidemiologists’ work has been impaired.

CONCLUSIONS AND RECOMMENDATIONS

It was agreed that “geographical analyses” refer to studies designed to exploit information on the spatial location in the data.

1. **When is it appropriate to use geographical analyses in public health decision-making?**

   Geographical analyses are appropriate when outcomes or exposures or a combination of both have a spatial structure. Studies of this nature can assist in public health decision-making.

   In particular, geographical analyses of the distribution of risk factors can be useful in prioritizing preventive measures. Disease mapping is useful for health service provision and targeting interventions if avoidable risk factors are known. Geographical studies of disease and environmental exposures may in some cases be sufficient by themselves to justify action, for example if the exposure-disease association is specific, the latency is short and the exposure is spatially defined.

2. **If appropriate, what is the methodology of choice?**

   No methodology of choice can be recommended in general. Analytical methods should be selected on the basis of the structure of the data to be analysed and of the hypotheses to be investigated (e.g. individual or aggregated data, presence of putative foci of risk). In most circumstances, however, it might be helpful to envisage a first level of descriptive analysis, to be followed by more specific, and problem-dependent, analyses involving parameter estimation and hypothesis testing. These will often be based on multivariate techniques and statistical modelling.

3. **Is a disease cluster, with no prior hypothesis, a sufficient cause for public health action?**

   The reporting of any disease cluster, even in the absence of a hypothesis defined a priori, should never be ignored but critically evaluated. The process of decision-making following cluster detection should be informed by considerations concerning the plausibility of any post-hoc hypothesis, its relevance in public health terms, the feasibility of possible preventive measures, and the resources needed for further investigation. While further evaluation is needed of the conditions under which public health action following cluster detection should be taken, it is noted that the simple statistical evidence of a localised excess is not sufficient to warrant intervention. Governments and other
agencies should proceed with caution when communicating the occurrence of possible disease clusters.

4. **Is it possible to depict a screening device for geographical clusters, in the context of public health surveillance?**

Public health surveillance, for example from population registries, may be valuable in responding to cluster reports. It is technically possible to devise tools that systematically screen for geographical clusters, allowing for the problem of multiple testing. However, for etiologic research, screening programmes are likely to be fruitful only in special circumstances, where highly specific exposure-disease associations and large attributable risks are involved. If surveillance is undertaken, it should be based on a clear protocol on the action to be taken if notable clusters are detected.

5. **Is it possible to use geographical methods to monitor public health interventions?**

In some circumstances geographical analyses have been used for monitoring the impact and the effectiveness of public health intervention. This might be more informative when dealing with public health measures with immediate or short-term effects.

6. **Are the new methods proposed of added value in generating hypotheses?**

Traditional methods for investigating spatial patterns of disease/exposure can be valuable in several circumstances to inform public health action, provided the data are valid, accurate and complete. However, recently developed methods of analysis, designed to deal with the spatial component of the data, have the potential to provide results that improve or correct those obtained using conventional methods, especially with small area or individual case location data. The presence of underlying geographical patterns can be more easily detected and their properties better described. In addition, such new methods have a greater potential in generating new hypotheses from geographically referenced health data.

7. **Do geographical analyses contribute to the strengthening of the evidence of the causal nature of an association?**

Geographic analyses with no information at the individual level are vulnerable to bias. However, while individually based epidemiological studies are in general needed to demonstrate the causal nature of an exposure-disease association, geographical analyses can help strengthen the available evidence. For some spatially distributed exposures such as environmental pollution, geographical studies are appropriate designs and can provide useful evidence in assessing causality, especially when the appropriate time scale is accounted for.

8. **How to communicate to decision-makers the results of geographical analyses?**

Providing public health decision-makers with results of geographical analyses, without addressing the underlying assumptions and discussing the implications, should be avoided. Communication of results should be based first on simple tabulations of the data and complemented by analyses addressing clearly specified hypotheses.

If maps are to be used, they should be accompanied by appropriate indices of uncertainty and variability along with some word of caution on interpretation. It is important that overall evaluations of the findings and conclusions be given.

9. **Is it possible to sketch a cost-utility evaluation of geographical methodologies in public health?**

Some technologies, such as geographic information systems (GIS), used for geographical analyses require data availability and can involve substantial financial investments. It is therefore appropriate to identify the direct benefits associated with this investment. These benefits include the capability of: undertaking rapid screening of apparent disease clusters to address the need for a ad-hoc study; bringing together the available information on geographically-based factors and/or confounders
potentially of relevance; targeting or delivering public health intervention or environmental health controls more efficiently; focussing attention or investigations on the areas where intervention might be most beneficial; contributing to a surveillance scheme when needed.

10. To what extent does data quality affect the different geographical methods?

As in all epidemiological studies, high data quality is crucial in geographical analyses. At the small area level, however, even relatively minor inconsistencies might have a large impact on the findings, especially with routinely collected data (e.g. population estimates derived from census data).
Annex 1

PARTICIPANTS

Temporary Advisers

Dr Ben Armstrong
Environmental Epidemiology Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom

Dr Olav Axelson
Department of Occupational Medicine, University Hospital, S-58185 Linkoping, Sweden

Dr Nikolaus Becker
Deutsches Krebsforschungszentrum Abteilung Epidemiologie, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

Dr Luisa Bernardinelli
Dipartimento di Scienze Sanitarie Applicate e Psicocomportamentali, Università di Pavia, Via Bassi 21, 27100 Pavia, Italy

Dr Julian Besag
Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195 USA

Dr Nicky Best
Department of Epidemiology & Public Health, Imperial College, School of Medicine at St Marys, Norfolk Place, London W2 1PG, United Kingdom

Dr Annibale Biggeri
Dipartimento di Statistica “G. Parenti”, Università di Firenze, Viale Morgagni, 5950134 Firenze, Italy

Dr John Francis Bithell
Lecturer in Statistics, Department of Statistics, 1 South Parks Road, Oxford OX1 3TG, United Kingdom

Dr Dankmar Böhning
Department of Epidemiology, Free University Berlin, Fabeck Str. 60-62, Haus 562, 14195 Berlin, Germany

Dr Cesare Cislaghi
Istituto di Biometria, Via Venezian 1, 20133 Milano, Italy

Dr Allan Clark
School of Informatics, University of Bell Street, Abertay, Dundee DD1 1HG, United Kingdom

Dr Noel A. C. Cressie
Department of Statistics, Iowa State University, Ames, Iowa IA 50011-1210, USA

Dr Fabio Divino
Dipartimento Statistico Universitario, Università di Firenze, Viale Morgagni, 59, 50134 Firenze, Italy

Dr Emanuela Dreassi
Dipartimento Statistico Universitario, Università di Firenze, Viale Morgagni, 5, 50134 Firenze, Italy
Dr Irene Schmidtmann  
Institut für Medizinische Statistik und Dokumentation, Klinikum der Johannes Gutenberg-Universität, 55101 Mainz, Germany

Dr Toshiro Tango  
Director, Division of Theoretical Epidemiology, The Institute of Public Health, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan

Dr Benedetto Terracini  
Dipartimento di Scienze Biomediche e Oncologia Umana, Università di Torino, Via Santena, 710126 Torino, Italy

Dr Jean-François Viel  
Department of Public Health, Faculty of Medicine, 2 Place Saint Jacques, 25030 Besançon, France

Dr Lance A Waller  
Division of Biostatistics, University of Minnesota, Box 303 Mayo Building, 420 Delaware St. SE, Minneapolis, MN 55455-0392, USA

Dr Stephen Walter  
McMaster University, Dept of Clinical Epidemiology and Biostatistics, HSC-2C16,1200 Main St W, Hamilton, Ontario L8N 3Z5, Canada

Dr Iris K. Zöllner  
Landesgesundheitsamt Baden Württemberg Referat Epidemiologie, Hoppenlaustrasse 7, 70174 Stuttgart, Germany

Observers

Dr Mario Braga  
Agenzia per i Servizi Sanitari Regionali, Palazzo Italia, Piano VIII, Piazza G. Marconi, 25, 00144 Roma, Italy

Dr Francesca Gallo  
Direzione Centrale delle Statistiche su Popolazione e Territorio, Istituto Nazionale di Statistica (ISTAT), Via Adolfo Ravà, 150, 00144 Roma, Italy

Dr Francesco Forastiere  
Dirigente, Unità Operative Epidemiologiche, Osservatorio Epidemiologico della Regione Lazio, Via di S. Costanza, 53, 00198 Roma, Italy

Dr Wolfgang Hoffmann  
Universität Bremen, Postfach 330 440, D28334 Bremen, Germany

Dr Corrado Magnani  
Servizio Epidemiologia Tumori, Università di Torino, Via Santena, 7, 10126 Torino, Italy

Dr Paola Michelozzi  
Osservatorio Epidemiologico della Regione Lazio, Via di Santa Costanza, 53, 00198 Roma, Italy

Dr Bart Ostro  
Air Pollution Epidemiology Unit, California Environmental Protection Agency, 2151 Berkley Way, Annex 11, Berkley, California 94704 USA
Dr Adele Seniori Costantini  
Occupational Epidemiology Branch, Centro per lo Studio per la Prevenzione Oncologia (CSPO), Via di San Salvi, 12, 50134 Firenze, Italy

World Health Organization

**International Agency for Research on Cancer**
Dr Marco Martuzzi  
Unit of Radiation and Cancer

**European Centre for Environment and Health**
Dr Roberto Bertollini  
Director, ECEH Rome Division

Dr Alexander Kuchuk  
Manager, EHIS, ECEH Bilthoven Division

Ms Candida Sansone  
Programme Assistant, ECEH Rome Division
Annex 2

WORKING PAPERS

ICP/EHH 001/1  Provisional list of working papers
ICP/EHH 001/2  Provisional agenda
ICP/EHH 001/3  Scope and purpose
ICP/EHH 001/4  Provisional programme
ICP/EHH 001/5  Provisional list of participants
ICP/EHH 001/6  Public health implications of ecological analyses, by Dr Olav Axelson, Sweden
ICP/EHH 001/7  Advances of etiologic knowledge have consequences on the scope of cancer mapping, by Dr Nikolaus Becker, Germany
ICP/EHH 001/8  Ecological regression with errors in covariates, by Dr Luisa Bernardinelli, Italy
ICP/EHH 001/9  Use of geographical data for public health decision-making in environmental health: a review of the European experience, by Dr Roberto Bertollini, WHO/ECEH Rome, Italy
ICP/EHH 001/10 Modelling risk from a disease in time and space, by Professor Julian Besag, USA
ICP/EHH 001/11 Bayesian ecological modelling, by Dr Nicky Best, United Kingdom
ICP/EHH 001/12 Non parametric modelling of spatial interaction function in geographical epidemiology, by Dr Fabio Divino, Italy; Dr Annibale Biggeri, Italy
ICP/EHH 001/13 Case control analysis of risk around putative sources, by Dr Annibale Biggeri, Italy
ICP/EHH 001/14 Exploratory methods in risk assessment, by Dr John Francis Bithell, United Kingdom
ICP/EHH 001/15 Disease mapping with hidden structures using mixture models, by Dr Peter Schlattman, Germany; Professor Dankmar Böhning, Germany
ICP/EHH 001/16 Air pollution and lung cancer: evidence from bio-indicators, by Dr Cesare Cislaghi, Italy
ICP/EHH 001/17 Small-area and point-level Bayesian models for inference on extremes in disease maps, by Dr Noel A. C. Cressie and Dr Hal S. Stern USA
ICP/EHH 001/18 Risk of congenital anomalies near hazardous waste landfill sites in Europe: The Eurohazcon Study, by Dr Helen Dolk, United Kingdom
ICP/EHH 001/19 Spatial regression models in epidemiological studies, by Dr Juan R. Ferrandiz, Spain
ICP/EHH 001/20 Geographical variation in health and environmental risk factors: maps, clusters, communication, by Dr Tony Fletcher, United Kingdom
ICP/EHH 001/21 Penalized pseudolikelihood inference in spatial interaction models with covariates, by Dr Fabio Divino, Italy, Dr Arnoldo Frigessi, Norway
ICP/EHH 001/22 Credibility: A new approach to disease clustering for uncertain locations, by Mr Geoffrey M. Jacquez, USA
ICP/EHH 001/23 Selected case studies in Bayesian disease mapping for health and health service research in Ireland, by Dr Alan Kelly, Ireland
<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICP/EHH 001/24</td>
<td>Statistical evaluation of disease cluster alarms, by Dr Martin Kulldorff, USA</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/25</td>
<td>Multilevel modelling of the geographical distributions of diseases, by Dr Ian H. Langford, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/26</td>
<td>Markov Chain Monte Carlo methods for putative sources of hazard and general clustering, by Dr Andrew B. Lawson, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/27</td>
<td>Disease mapping and its uses, by Dr Andrew B. Lawson</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/28</td>
<td>Addressing multiple goals in evaluating region-specific risk using Bayesian methods, by Erin M. Conlon, MS and Dr Thomas A. Louis, USA.</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/29</td>
<td>Heterogeneity of disease rates, by Dr Marco Martuzzi, France and Dr Michael Hills, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/30</td>
<td>Sequential monitoring of low event rates: an application in environmental epidemiology, by Dr Armand Maul, France</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/31</td>
<td>Bayesian mapping of disease, by Dr Annie Mollié, France</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/32</td>
<td>The use and misuse of geographical information systems in epidemiology, by Dr Raymond Neutra, USA</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/33</td>
<td>Lung cancer near point emission sources, by Dr Goran Pershagen, Sweden</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/34</td>
<td>Identifying clusters in disease maps, by Dr Irene Schmidtmann, Germany</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/35</td>
<td>Comparison of general tests for spatial clustering, by Dr Toshiro Tango, Japan</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/36</td>
<td>Pleural cancer in small areas as a marker of environmental and occupational exposure to asbestos, by Professor Benedetto Terracini, Italy</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/37</td>
<td>Environmental epidemiology, public health advocacy and policy, by Dr Jean-Francois Viel, France</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/38</td>
<td>The power of focused score tests under mis-specified cluster models, by Dr Lance A Waller, USA</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/39</td>
<td>An analysis of determinants of regional variation in cancer incidence: Ontario, Canada, by Dr Stephen Walter, Canada</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/40</td>
<td>Different cluster tests in application to German cancer maps, by Dr Iris K. Zöllner, Germany</td>
<td></td>
</tr>
<tr>
<td>ICP/EHH 001/41</td>
<td>Edge effects in disease mapping, by Dr Andrew B. Lawson, United Kingdom; Dr Emanuela Dreassi, Italy; Dr Annibale Biggeri, Italy</td>
<td></td>
</tr>
</tbody>
</table>