Small-area spatio-temporal analyses of participation rates in the mammography screening program in the city of Dortmund (NW Germany)

Dorothea Lemke1,2*, Shoma Berkemeyer3, Volkmar Mattauch4, Oliver Heidinger4, Edzer Pebesma2 and Hans-Werner Hense1,4

Abstract

Background: The population-based mammography screening program (MSP) was implemented by the end of 2005 in Germany, and all women between 50 and 69 years are actively invited to a free biennial screening examination. However, despite the expected benefits, the overall participation rates range only between 50 and 55%. There is also increasing evidence that belonging to a vulnerable population, such as ethnic minorities or low income groups, is associated with a decreased likelihood of participating in screening programs. This study aimed to analyze in more detail the intra-urban variation of MSP uptake at the neighborhood level (i.e. statistical districts) for the city of Dortmund in northwest Germany and to identify demographic and socioeconomic risk factors that contribute to non-response to screening invitations.

Methods: The numbers of participants by statistical district were aggregated over the three periods 2007/2008, 2009/2010, and 2011/2012. Participation rates were calculated as numbers of participants per female resident population averaged over each 2-year period. Bayesian hierarchical spatial models extended with a temporal and spatio-temporal interaction effect were used to analyze the participation rates applying integrated nested Laplace approximations (INLA). The model included explanatory covariates taken from the atlas of social structure of Dortmund.

Results: Generally, participation rates rose for all districts over the time periods. However, participation was persistently lowest in the inner city of Dortmund. Multivariable regression analysis showed that migrant status and long-term unemployment were associated with significant increases of non-attendance in the MSP.

Conclusion: Low income groups and immigrant populations are clustered in the inner city of Dortmund and the observed spatial pattern of persistently low participation in the city center is likely linked to the underlying socioeconomic gradient. This corresponds with the findings of the ecological regression analysis manifesting socioeconomically deprived neighborhoods as risk factors for low attendance in the MSP. Spatio-temporal surveillance of participation in cancer screening programs may be used to identify spatial inequalities in screening uptake and plan spatially focused interventions.

Keywords: Mammography screening, Participation rates, Spatio-temporal modelling, INLA
Background
The implementation of a nation-wide, population-based mammography screening program (MSP) started in Germany by the end of the year 2005. The stepped implementation process was completed in the state of North-Rhine Westphalia in 2009. All resident women aged between 50 and 69 years are actively invited to a mammography screening examination every two years. The participation is voluntary and free of cost. Mammography screening is a procedure of secondary cancer prevention with an aim of detecting breast cancer in early stages where therapy is less invasive (e.g., breast-conserving therapies instead of mastectomy), remaining lifespans are extended and, ideally, breast cancer mortality is reduced. Despite these expected benefits and the free provision by all statutory health insurances, the overall participation rates range only between 50 and 55 % [1]. Population-based surveys demonstrate that significant gaps exist in screening mammography uptake across population subgroups [2]. These differences are believed to substantially contribute to a higher prevalence of late stage breast cancer at diagnosis among vulnerable populations, including racial and ethnic minorities [3] and low-income groups [4, 5]. More specifically, living in an economically deprived neighborhood showed a decreased likelihood of participating in cancer screening programs and an increased risk of a late-stage breast cancer diagnosis with the correspondent unfavorable prognosis [2, 6–9]. To date, few studies analyzed the intra-urban variation of participation rates in mammography screening programs [2, 10] and to our knowledge none investigated the situation in Germany. We suggest that small-area analyses may provide important insights into the processes and factors that are associated with low participation rates and that this may help to develop spatially focused approaches to improve the participation rates in disadvantaged neighborhoods.

Therefore, this study aimed to investigate the spatio-temporal distribution of the participation rates in the mammography screening program at the neighborhood-level (e.g. statistical districts) of a large city in Germany and to identify important demographic and socioeconomic factors that influence the non-attendance to screening invitations.

Methods
Study region
Dortmund is a city in the federal state of North Rhine-Westphalia in northwestern Germany with a total population of 575,944 inhabitants in 2013. It is the largest city by area and population in the Ruhr district, a metropolitan area with some 5.1 million inhabitants which is the largest urban and industrial agglomeration in Germany. Dortmund is divided into 62 statistical districts with a median female population of 4086 inhabitants per statistical district. The city’s population is characterized by a high proportion of immigrants from southeast Europe and Turkey. The coal crisis in the end of 1970 led to a massive reduction of jobs in the coal and steel industry which resulted in high unemployment rates to this day. Immigrants and their descendants grew up in more socially deprived neighborhoods than many of the autochthonous population. This resulted in a strong spatial segregation of populations with a migration background and with low economic status within the city [11].

Participation rates and geo-referencing
The statistical districts were used as the geographical reference system in which the MSP participation rates were assessed. The residence addresses of all MSP participants for the years 2005 to 2013 were stored at KV.it Dortmund, the institution which administers the MSP documentation software MaSc [12]. KV.it assigned MSP participants to one of the 62 statistical districts by linking their home addresses to a comprehensive list of street addresses for each district. A list of individual, anonymized participants who were geo-referenced to one of the statistical districts was then transferred to the Institute of Epidemiology and Social Medicine at the University Münster [13] where all subsequent analyses were carried out.

The years 2005 and 2006 were excluded from the present analyses to avoid contamination with the various organizational aspects of the stepped-up implementation of the MSP. All eligible women receive a biennial invitation to the screening program, hence, we chose to analyze three 2-year periods: 2007/2008, 2009/2010, and 2011/2012. The participation rates were calculated using the aggregated numbers of participants and the averaged female background population (age group 50–69) for each two year-period.

Spatio-temporal mapping and regression
The spatio-temporal distribution of participation rates was analyzed within a hierarchical Bayesian framework using a multivariate binomial regression model (spatio-temporal odds model): Let \( n_{it} \) denote the number of eligible women resident in district \( i \) and period \( t \) and \( Y_{it} \) the number of participants in breast cancer screening, with \( i = 1, \ldots, 62 \) and \( t = 1, 2, 3 \). We assumed that the observed number of participants \( Y_{it} \) had a binomial distribution with parameters \( n_{it} \) and \( \theta_{it} \) (probability of participation). At a second level, the probability of participation \( \theta_{it} \) was then decomposed on the logit scale into an overall participation rate \( \alpha \), main spatial effects \( (u_i \text{ and } v_t) \) (constant in time), main temporal effects
(unstructured ($\Phi_t$) and structured ($\gamma_t$)), and a space-time interaction term ($\psi_{it}$).

$$Y_{it} \sim \text{Binomial}(n_{it}, \theta_{it})$$  \hspace{1cm} (1)

$$\logit \theta_{it} = \alpha + u_i + v_t + \Phi_t + \gamma_t + \psi_{it}$$  \hspace{1cm} (2)

The proposed space-time models, assuming a non-parametric time trend and a spatio-temporal interaction term, were introduced by Knorr-Held [14] and are an extension of the spatial model introduced by Besag et al. [15]. All model terms were treated as random variables: The spatially unstructured random effect ($u_i$) was considered independent and identically distributed (iid) with zero mean and unknown precision ($\tau_u$). To account for the assumption of correlated participation rates in nearby statistical districts, the spatially structured effect ($v_t$) is modelled for each 62 districts as an intrinsic Gaussian Markov random field with unknown precision ($\tau_v$). This specification is also called a conditionally auto-regressive (CAR) prior and was introduced by Besag et al.[15]. In order to insure the identifiability of the intercept $\alpha$ (overall participation rate), a sum-to-zero constraint was imposed on the $v_t$'s [16]. The unstructured temporal effect ($\Phi_t$) was also modelled iid with zero mean and unknown precision. For the structured time effect ($\gamma_t$) random walks of first order were considered [17, 18]. The interaction term ($\psi_{it}$) can be specified in several ways [14], here it is assumed that the two unstructured effects ($v_t$ and $\gamma_t$) interact [17, 18]. Therefore, the interaction effect was also specified as zero mean normal with unknown precision (iid, i.e. $\psi_{it} \sim N(0, \tau_{\psi})$). The distribution of the hyperpriors was specified as follows: Minimally informative priors were specified on the log of the unstructured effect precision ($\log \tau_v \sim \text{logGamma}(1,0.001)$) and on the log of the structured effect precision ($\log \tau_u \sim \text{logGamma}(1,0.001)$). For the unstructured time effect, a $\log \tau_u \sim \text{logGamma}(1,0.01)$ hyperprior was chosen. For the structured time effect and the interaction term, minimally informative priors (the default priors): $\log \tau_{\psi} \sim \text{logGamma}(1,0.00005)$ have been used. Altogether, the distribution of the hyperpriors resembles the ones used by Ugarte et al [18].

$$\logit \theta_{it} = \alpha + u_i + v_t + \Phi_t + \gamma_t + \psi_{it} + \beta x_{it}^T$$  \hspace{1cm} (3)

The specified model (Equation 3) was extended to $\beta x_{it}^T_p$, where $x_{it}^T$ contains the covariates with a space-time index in order to investigate potential risk factors associated with spatio-temporal variations in the participation rates. The covariates were taken from the atlas of social structure (Sozialstrukturalatlas) of Dortmund which is a collection of administratively collected data reflecting social inequalities and differences in the population [19]. These are grouped into the dimensions employment status, demography, income, welfare, and housing. A full description of the explanatory variables is given in Table 1. In order to account for multicollinearity, an initial correlation matrix was examined for high correlations among the variables. Variables with a high correlation (>0.8) were excluded from further regression analyses. For the three 2-year time periods, the data of 2008, 2010 and 2012, respectively, were included in the model, and all variables were dichotomized according to their median value. Following the suggestions of Rothman et al. [20-21] each covariate was fitted separately and model fit was assessed using the changes in deviance information criterion (DIC) (smaller values of DIC indicating more explained variance and better fit). A multivariable model was fit by selectively including variables, starting with those that showed the lowest DIC in univariable analyses, until the DIC could be no further reduced. For the Bayesian inference, the integrated nested Laplace approximation (INLA) approach was used as introduced by Rue et al. [22] and implemented in the R package R-INLA [21, 23, 24]. The Bayesian inference was also used to report the resulting odd ratios (OR) as point estimate (posterior mean) and 95% credibility intervals (CI) as a quantification of parameter uncertainty. All computations and visualizations were done in R v. 3.0.2 [25].

Ethics statement
KV.it [12] administrates the MSP documentation including the storage and management of the MSP participants consistent with the existing data protection legislation. For this study, KV.it aggregated participants in statistical districts so that individual women could be not identified. Data were transferred to the investigators in an anonymized form. Use of anonymized data for research purposes does not require a vote by ethics committee or an institutional review board.

Results
Mapping participation rates
The observed annual participation rates showed the overall biennial pattern of participation, i.e., one year with high number of participants and the subsequent year with lower participants (Fig. 1). Despite rising overall MSP participation rates for the three periods from 48% (2007/08) over 50% (2009/10) to 54% (2011/12) (Referenzzentrum MS), a concentration of statistical districts with low participation rates persisted in the inner city of Dortmund, while the outer districts had consistently higher participation rates (Fig. 2a–c). The modeled time trends in Fig. 2d demonstrate these increasing participation rates over the three periods, where the structured time effect ($\gamma_t$) was more pronounced than the unstructured time effect ($\Phi_t$). The spatial trends combine structured and unstructured effects (Fig. 2e) and
confirm, after accounting for covariates, that lower participation rates cluster in the inner city. Finally, the interaction analyses reveals a clear space-time pattern (Fig. 2f-h) which indicates that in 2007/8, on a generally low participation level, the participation rates were particularly low in the eastern districts. This changed in 2008/9 as lower participation persisted in the western and central parts of the city, while in 2010/11 a low participation rate was found in only one inner city district.

**Regression analyses**

Due to high correlations and close content relations, the variables: unemployment rate (total), persons with migrational background, and basic social welfare rate were excluded from the further analyses. The results of univariate and multivariable spatio-temporal regression analyses are summarized in Table 2. The odds ratios of the univariate spatio-temporal regression analyses demonstrate clearly that districts with high proportions of unemployed migrants or long-term unemployed residents showed a statistically significant impact on lower participation rates. In contrast, higher proportion of elderly population showed positive association with the participation rates. The other ecological variables clearly contained null values in their credibility intervals and were therefore considered as factors without relevant influence. In the multivariable analyses that simultaneously adjusted for spatio-temporal variation, the negative association of the proportion of unemployed migrants and long-term unemployed remained statistically significant.

| Dimension/Indicator | Variable                                                        | Median value (2008; 2010; 2012) | Definition                                                                 |
|---------------------|-----------------------------------------------------------------|-------------------------------|---------------------------------------------------------------------------|
| Employment/Unemployment | Employment rate [%]                                             | 48.2; 48.4; 50.1              | Proportions of regular employees of the employable population (>15 to <65 years) with primary residence in Dortmund. |
|                      | Employment trend [%]                                            | 3.5; 5.9; 7                   | Number of regular employees with primary residence in Dortmund in a five-year trend/comparison in percent. |
|                      | Unemployment rate (foreigner) [%]                              | 12.6; 12.5; 12.7             | Proportion of unemployed, foreign persons of the employable, foreign population (>15 to <65 years) in percent. |
|                      | Unemployment rate (<25 aged) [%]                               | 4.2; 4.5; 4.3                 | Proportion of unemployed young persons (15 to < 25 years) of the population in the same age group. |
|                      | Unemployment rate (long-term) [%]                              | 44.8; 41.3; 44.7             | Proportion of long-term unemployed persons (>12 months) to all unemployed persons in percent. |
| Demography           | Female population [%]                                          | 51.7; 51.5; 51.5             | Proportion of women to all inhabitants with primary residence in Dortmund. |
|                      | Foreign residents [%]                                           | 7.5; 7.1; 7.7                 | Proportion of persons without German nationality to all inhabitants with primary residence in Dortmund. |
|                      | Youth quotient                                                  | 20.5; 20; 19                  | Number of persons aged < 15 years per 100 persons aged 15 to < 65 years. |
|                      | Quotient of elderly                                             | 32.7; 32.7; 31.6             | Number of persons aged > 65 years per 100 persons aged 15 to < 65 years. |
|                      | Birth rate                                                      | 8; 7.4; 7.5                   | Number of births per 1000 inhabitants with primary residence in Dortmund by December 31. of each year. |
|                      | Mortality rate                                                  | 10.4; 10.5; 10.5             | Number of deaths per 1000 inhabitants with primary residence in Dortmund by December 31. of each year. |
| Social affairs        | Long-term social welfare (“Hartz IV”) rate [%]                  | 10.3; 11.9; 11.6             | Proportions of unemployed (>12 months) persons (<65 years) with demand for basic financial benefits (Arbeitslosengeld II) of the residential population. |
| Living/habitation     | Social housing rate [%]                                         | 8; 5.4; 6.2                   | Proportion of council-sponsored apartments on all rental appartements in percent. |
|                      | Living space per person                                        | 39.6; 40.2; 40.4             | The total living space in [m²] divided by the number of inhabitants with primary and secondary residence in Dortmund. |
Proportions of migrant unemployment and long-term unemployed above the median were associated with a significant 6 % and, 3 %, respectively, increase in risk for non-attendance in the mammography screening program.

Discussion

The present study analyzed small-area, intra urban variations of participation rates of the MSP in the city of Dortmund over three 2-year periods. An overall increase in the participation rates was observed over the study period, while the increase was unevenly distributed across the study area. There was a spatial concentration of statistical districts, mainly in the city center, with persistently low participation rates. Dortmund is known to have a strong gradient of socio-economical segregation [26]. Its population is characterized by a high proportion of residents which are likely to have a lower socio-economic status, because of the massive loss of jobs in mining and steel industries in this area after 1970 (also known as structural crisis) which was compounded in population segments with a migration background [11]. Muller and Berger [26] reported, in their investigation
Fig. 2 (See legend on next page.)
about neighborhood deprivation and prevalence of type 2 diabetes in Dortmund, that the inner city and parts of the western city are characterized by the highest level of socio-economic constraints including high proportions of immigrants, unemployed residents, residents with basic social welfare as well as a high population density and low level of incomes.

Therefore, it seems plausible that the observed spatial pattern in the participation rates is linked to the underlying socio-economic gradient. Despite the high availability of screening facilities in the inner city (Fig. 2a–c), it should be noted that these statistical districts had persistently low participation rates. In contrast, reduced participation rates in the southernmost districts may be attributable to a higher proportion of women with private health insurances who tend to abstain from public health offers. The highest rates of participation were found in the eastern parts of the city, that has more affluent statistical districts and screening facilities more nearby. The ecological regression analyses confirmed the spatio-temporal results by revealing that characteristics of disadvantage in statistical districts were related to an increased probability of non-participation in the MSP.

The identification of socio-economic risk factors at area-level as explanatory variables of non-attendance in mammography screening have been examined in previous studies [27–30] which also found that neighborhood income was an important determinant of participation [31–33]. Additionally, Peek and Han [34] reported that vulnerable groups such as the poor, the elderly, and minorities were often unaware of mammography screening programs and had a reduced awareness and a lack of information of disease prevention, diagnosis, and treatment [35]. Awareness of the program is unlikely to play a major role in Germany as all resident women were personally invited as part of a structured systematic program of early breast cancer detection; attitude towards disease prevention seems to be a more likely reason for the lower MSP attendance.

Regarding the overall increase in the participation rates over time, it should be kept in mind that the MSP is a rather recent program as compared to other European countries [36]. It became operational as part of routine care only by the end of the year 2005, and comprehensive implementation in North Rhine–Westphalia was not completed before the end of 2009 [37]. Therefore, the

Table 2 Estimation results of univariable and multivariable analysis using the spatio-temporal model

| Variable                          | pD  | DIC    |
|----------------------------------|-----|--------|
| Space-time CAR without covariates| 95.94 | 1707.09 |
| Explanatory variables            |     |        |
| Unemployment rate (foreigner) [%] | 0.95 | (0.91; 0.99) | 1706.279 | 0.94 | (0.90; 0.98) | 1706.56 |
| Living space per person          | 1.03 | (0.95; 1.13) | 1707.15 | 1.01 | (0.92; 1.12) |
| Unemployment rate (long-term) [%] | 0.98 | (0.94; 1.0) | 1707.174 | 0.97 | (0.94; 1.0) |
| Employment trend [%]             | 0.98 | (0.95; 1.01) | 1707.211 |
| Mortality rate                   | 1.02 | (0.98; 1.06) | 1707.305 |
| Female population [%]            | 1.03 | (0.98; 1.09) | 1707.432 |
| Social housing rate [%]          | 1.00 | (0.96; 1.05) | 1708.321 |
| Foreign residents [%]            | 0.95 | (0.89; 1.02) | 1708.38 |
| Quotient of elderly              | 1.07 | (0.99; 1.15) | 1708.561 |
| Basic social welfare rate [%]    | 0.97 | (0.91; 1.04) | 1708.566 |
| Unemployment rate (<25 aged) [%] | 0.99 | (0.94; 1.04) | 1708.711 |
| Youth quotient                   | 0.99 | (0.95; 1.04) | 1708.801 |
| Employment rate [%]              | 1.02 | (0.98; 1.07) | 1708.951 |
| Birth rate                       | 0.97 | (0.94; 1.01) | 1709.389 |

Estimates are odds ratios with their associated 95 % Credible Intervals (CI). Values below one denoting a reduced odd of participating in screening, and vice versa.
increasing participation rates over the study period may be mainly attributed to an increased efficiency of operational routines within the screening units which allowed for a growing numbers of screened women and it is probable that these observed spatial effects - especially within the first two study periods - are influenced by these technical and structural developments [38]. The spatio-temporal interaction effect adds to the spatial and temporal findings in that it identifies districts where the observed participation rates were reduced as compared to the entire city and throughout the study period [39]. Holding the spatial component constant confirmed the increasing overall trend of the participation rates, while holding the temporal trend constant confirmed that statistical districts in the city center had consistently reduced attendance rate in mammography screening.

This study has several strengths and limitations. An obvious strength has been the use of the Bayesian hierarchical framework in order to borrow strength from spatial and temporal neighbors to reduce the high variability inherent in the estimators, in particular, when numbers (disease counts and/or background population) are unstable [18, 39, 40]. Also, the inclusion of a space-time interaction effect is an added strength of this study, because the participation rates of mammography screening may be plausibly assumed to be dependent in space and time. Adjustment for the spatial, temporal, and space-time interaction effects depicts more clearly how the spatial pattern of the participation rates evolved over time, while the intersection of space and time is seldom considered to disentangle the complex determinants of health-related behavior and diseases [39]. Furthermore, the use of a non-parametric time trend has been a more plausible assumption than a linear time trend, because not all census tracts showed a linear increase or decrease in their participation rates. However, the analysis of effects in our study was confined to only three time periods, and hence requires caution in interpretation. The use of integrated nested Laplace approximations (INLA) reduced computing time substantially while attaining a high degree of accuracy, when fitting large, complex data sets at detailed geographic levels as used in spatio-temporal disease mapping. Given the inherently ecological nature of this study, the parameter estimates may not be used for making inferences on the individual level and therefore must not be interpreted causally. However, the results provide important hints on how social, cultural, and contextual factors may influence the attendance in mammography screening. Thus, despite the ecological nature of our study, the results may be used to provide spatially focused interventions to improve participation in disadvantaged city districts. Another limitation results from the fact that the precise number of invited women was not available due to data privacy regulations in Germany. The denominator used for calculating the participation rates contained therefore the whole female population in that age group which also included women non-eligible to screening (e.g., because of prevalent breast cancer). Thus, the participation rates may be slightly underestimated but this is not perceived as biasing the spatial associations. Another limitation results from the aggregation of the number of participants to a period of 24 month. Because the invitations to biennial screening were continuously mailed throughout the 22–26 month period, a certain amount of misclassifications is to be expected. However, as the general trend of the participation rates showed a clear biennial pattern it seems safe to assume that the main spatio-temporal process of the participation rates was captured with the temporal aggregation employed in this study.

Conclusions
This study analyzed the intra-urban participation rates of a mammography screening program within a hierarchical Bayesian framework using spatio-temporal disease models to identify regions and risk factors of low attendance. Despite a general temporal trend with increasing participation rates, spatial clustering of persistently lower participation rates was observed in the inner city districts, which are known as the socio-economically most deprived neighborhoods of Dortmund. This corresponds with the findings of the ecological regression analysis manifesting indicators of socio-economic constraint in a neighborhood as risk factors for low attendance in the MSP. The spatio-temporal interaction effect showed that the participation rates developed spatially unequally over time and that certain districts had low participation rates throughout the study period. Spatio-temporal surveillance of the participation rates and focused intervention could help identifying and reducing spatial inequalities in the uptake of mammography screening.

Competing interests
The authors declare that they have no competing interest.

Authors' contributions
DL and HWH designed the study. VM and OH provided the data. DL carried out all analyses and drafted the first version of the manuscript. SB carried out confirmatory checks. DL, SB, EP and HWH interpreted the results. All authors were involved in critical review of the manuscript and assented to final version.

Acknowledgements
We would like thank KV.it (Dortmund) for georeferencing the screening participants. Additionally, we like to thank the statistical department of Dortmund for supply of the geometrical boundary information, the background population, and the covariate data.

Author details
1Institute of Epidemiology and Social Medicine, Medical Faculty, Westfälische Wilhelms-Universität Münster, Albert-Schweitzer-Campus 1 D3, D 48149 Münster, Germany. 2Institute for Geoinformatics, Geosciences Faculty, Westfälische Wilhelms-Universität Münster, Münster, Germany. 3Reference Center for the Mammography Screening Program, University Hospital, Westfälische Wilhelms-Universität Münster, Münster, Germany. 4Epidemiological Cancer Registry North Rhine-Westphalia, Münster, Germany.
