

Supplementary materials

Neighbourhood risk index

The neighbourhood risk index is a relative measure of multiple risk domains at the level of the neighbourhood in the Netherlands. The model of multiple risk, which underpins the index, is based on the idea of distinct dimensions of risk which can be recognised and measured separately (Glymour, Avendano, & Kawachi, 2014; Wilson, 1987; see also Noble, McLennan et al., 2019).

The neighbourhood risk index is a measure of multiple risk based on combining together distinct domains of risk, which consist of single or multiple indicators. These domains were based on the English Indices of Deprivation 2019 (Noble et al., 2019) and on the work of Leventhal and Brooks-Gunn (2000). The ten domains include the following:

- Income
- Employment
- Education
- Health
- Crime
- Housing
- Obstacles to services
- Green space
- Address density
- Ethnicity

Constructing the neighbourhood risk index

Data time point and geography scale

As far as possible, each indicator was based on data from the most recent time point available. This means that there is not a single consistent time point for all indicators and/or domains. The neighbourhood risk index is linked to the 6-digit zip code (PC6) in the Netherlands. As data on the PC6 had large numbers of missings, we used data from 5-digit zip code (PC5)-areas. Alternatively, when this data was not available, we used data on the neighbourhood-area level. The PC6 includes four

numbers and two letters (e.g., 1234AB), the PC5 one letter less. The first two numbers indicate a region, the second two numbers a city, village, or neighbourhood, and in general the four numbers indicate areas within the same municipality. The letters are more specific and indicate (a part of) a street. Municipalities consist of multiple neighbourhoods (CBS, 2021). In general data was used from PC5-areas¹ 2017, 2018, and 2019 and from neighbourhood areas² 2015, 2018, and 2019.

The domains and indicators

Ten domains of deprivation are combined to produce the overall neighbourhood risk index, and each domain contains a single or a number of component indicators. The criteria for inclusion of these indicators are that they should be 'domain specific', appropriate for the purpose of measuring risk, and available for the whole of the Netherlands at a PC5 or neighbourhood level in a consistent form. Further details about the purpose of each domain and the use of indicators are described below.

Income risk domain

The income risk domain aims to measure the proportion of the population in an area experiencing risk relating to income. Based on previous studies 'low income', compared to 'high income', indicated risk (e.g., Chaudry & Wimer, 2016; McBride Murry et al., 2011; Runarsdottir & Vilhjalmsón, 2019). The used definition of low income includes those people that are out of work and those that are in work but

¹ Detailed information is available through this link: <https://www.cbs.nl/-/media/cbs/dossiers/nederland-regionaal/postcode/statistische-gegevens-per-vierkant-en-postcode-2017.pdf>. The data was bought from CBS for research purposes. Data is not publicly available and may not be provided to third parties. Data is available to third parties in an edited, non-reducible form (e.g., in the form of the risk index).

² Detailed information and data is publicly available and accessible through <https://www.cbs.nl/nl-nl/maatwerk/2015/48/kerncijfers-wijken-en-buurtten-2015>; <https://www.cbs.nl/nl-nl/maatwerk/2018/30/kerncijfers-wijken-en-buurtten-2018>; and <https://www.cbs.nl/nl-nl/maatwerk/2019/31/kerncijfers-wijken-en-buurtten-2019>

have low earnings. The best available data was the number of residents with social security benefits related to unemployment, social assistance, and/or incapacity for work. This was data on the PC5 level over the year 2018. To use the data as an indicator for current purposes, we calculated the percentage of residents with social security benefits in an area compared to the total residents in an area. A higher score indicated more risk.

Employment risk domain

The employment risk domain measures the proportion of the working age population that is excluded from the labour market (e.g., Noble, McLennan et al., 2019). This includes people who would like to work but are unable to do so due to unemployment, caring responsibilities, disability, or sickness. The best available data was the percentage of the employed labour force in the population compared to the total population (working and non-working population). This was data on the neighbourhood level over the year 2019. To use the data as an indicator for current purposes, we used the formula: $100 - \text{percentage of employed labour force}$. In this way, a higher score indicated risk.

Education risk domain

The education risk domain aims to measure the education level in the local population. Based on previous studies 'low education level', compared to 'high education level', indicated risk (e.g., Friedli, 2009; Saja, Teo, Goonetilleke, & Ziyath, 2018). The best available data was the number of people (aged 15 to 75) with a low education level, a medium education level, and a high education level. Low, medium, or high education level were according to the Dutch classification of education program levels (CBS, 2022). This was data on the neighbourhood level over the year 2019. To use the data as an indicator for current purposes, first the number of people with low, medium, and high education level were added. Then, percentages of low and highly educated people were calculated. The percentage was

not calculated for secondary education level as this is average in the Netherlands. For the indicator, we used the formula: $100 - \text{percentage of highly educated people} + \text{percentage of low educated people}$. The index score ranged from 0-200. Based on this, if an area only contained people with an average education level, the index score would be 100. A higher score indicated more risk (more low educated people).

Health risk domain

The health risk domain aims to measure the impairment of quality of life through poor physical or mental health. The best available data was the percentage of youth (until 23 years old) who received youth services. This concerns a) youth assistance: help and care for young people and their parents with psychological, psychosocial, and/or behavioural problems, an intellectual disability of the youth, or parenting problems. As well as, b) youth protection: a child or youth is 'placed under supervision' or 'placed under guardianship' in order to eliminate the threat to the safety and development of the child, and c) help and control for children 12 years and older, who have been in contact with the police before their 18th birthday and have received an official police report. This was data on the neighbourhood level over the year 2019. This data could be used as an indicator in its current form. A higher score indicated more risk.

Crime domain

The crime domain measures the risk of different kinds of crime at a local level. Previous studies showed that more crime, compared to less crime, indicated risk (e.g., Aneshensel & Sucoff, 1996; Noble et al., 2019). The best available data was the total amount of theft, vandalism or offense against public order, and violent or sexual crimes per 1000 residents. This was data on the neighbourhood level over the year 2018. To use the data as an indicator for current purposes, we added the three types of crimes together. A higher score indicated more risk.

Housing risk domain

The housing risk domain measures the number of home-owners in a local area. Based on previous research less home-owners, compared to more home-owners and/or renters, indicated more risk (Badland, Foster, Bentley et al., 2017; Haurin, Dietz, & Weinberg, 2002). The best available data was the percentage of owner-occupied homes and rental homes. This was data based on the PC5 level over the year 2017. To use the data as an indicator for current purposes, we used the formula: $100 - \text{percentage of owner-occupied homes} + \text{percentage of rental homes}$. A higher score indicated more risk (more rental homes).

Obstacles to services domain

This domain measures the physical accessibility and proximity of local services. Based on the English index (Noble et al., 2019), it is riskier when a service is more distant. The best available data was the average distance in kilometres to the closest supermarket, general practice, day care, and primary school. This was data on the PC5 level over the year 2017. To use the data as an indicator for current purposes, we added the distances of the four local services together. A higher score indicated more risk.

Green space risk domain

The green space domain measures the amount of greenery of the local environment. Based on previous studies less green space, compared to more green space, indicated risk (e.g., Abraham, Sommerhalder, & Abel, 2010; Houlden, Porto de Albuquerque, Weich, et al., 2021; Li, Deal, Zhou et al., 2018). The best available data was total area of the Netherlands, total recreation area (which includes areas such as parks, sport fields, allotment gardens, zoos, amusement parks, and campsites), and total forest and open natural area (which includes areas such as forests, heathlands, and beaches). This was data on the neighbourhood level over the year 2015. To use the data as an indicator for current purposes, we added the areas of recreation area and

forest and open natural area together. Then we calculated the percentage of green space out of the total area of the Netherlands and used the formula: $100 - \text{percentage of green space}$. In this way, a higher score indicated risk.

Address density domain

This domain measures the density of addresses per local area. Based on previous studies a higher density, compared to a lower density, indicated risk (Vollebergh, Van Dorsselaer, Monshouwer et al., 2006; Zijlema, Klijs, Stolk et al., 2015). The best available data was the average number of addresses per square kilometre. This was data on the PC5 level over the year 2018. This data could be used as an indicator in its current form. A higher score indicated more risk.

Ethnicity domain

The ethnicity domain aims to measure the ethnic diversity in a local area. Previous studies showed that larger percentages of foreign-born residents indicated risk (e.g., see Leventhal, & Brooks-Gunn, 2000). The best available data was the percentage of people with Dutch ethnicity (people whose parents were born in the Netherlands; versus people whose parent(s) was/were not born in the Netherlands). This was data on the PC5 level over the year 2019. To use the data as an indicator for current purposes, we used the formula: $100 - \text{percentage of people with Dutch ethnicity}$. In this way, a higher score indicated risk.

Combining the domains

Each domain was constructed separately from the component indicator(s). Then, based on previous work by Theall and colleagues (2012), these raw scores on the domains were standardized with a mean of 0 and a standard deviation of 1, categorized so that 1 standard deviation above the mean meant "at risk". This resulted in z-scores that indicated whether or not a specific area was at risk for a certain domain. Next, the ten domain scores (0 = no risk; 1 = risk) were added to

create a total index score. A total risk index was only calculated for those areas that had a score on more than half (6 or more) of the domains. For areas with six or more missing sub-index scores, the total risk index was considered missing. For the areas with one to five missing domain scores, the total index score was multiplied by $(10/(10 - \text{number of missing domain scores}))$. The total risk index scores ranged from 0-10. Following the work from Theall and colleagues (2012), the cumulative risk index was categorized as low (0 risk factors), moderate (1-2 risk factors) or high (3-4 risk factors), or extremely high (>4 risk factors) risk.

Usage

The neighbourhood risk index, and each of the domains, can be used to rank every neighbourhood in the Netherlands according to the risk on the level of the neighbourhood experienced by the people living there. This data can be used to determine the risk of an area compared to other areas in the Netherlands.

Researchers can be contacted for the data for the overall neighbourhood risk index and each of the domains separately.

Points to consider when using the data

Points to consider are similar from the English Index of Deprivation (Noble et al., 2019).

- The neighbourhood-level domain scores and index provide a description of areas. However, this description does not apply to every person living in those areas. Many deprived people live in not-at-risk areas, and many non-deprived people live in at-risk areas.
- Those areas that are not identified as relatively at-risk by the neighbourhood-level domains and/or index are not necessarily affluent or wealthy areas. So, a lower ranked area could be described as less at-risk, but not as richer or more affluent. For example, the indicator for the income risk domain was the number of residents with social security benefits related to unemployment,

social assistance, and/or incapacity for work. A low number of people with social security benefits does not necessarily mean that there is a high number of people with high incomes.