Title: Social clustering of unvaccinated children in schools in the Netherlands

Authors: Don Klinkenberg1, Albert Jan van Hoek1, Irene Veldhuijzen1, Susan Hahné1, Jacco Wallinga1,2

Affiliations

1. Centre for Infectious Disease Control, National Institute for Public Health and the Environment, P.O.Box 1, 3720 BA Bilthoven, The Netherlands
2. Department of Medical Statistics, Leiden University Medical Centre, P.O. Box 9600, 2300 RC Leiden, The Netherlands

* Corresponding Author: don.klinkenberg@rivm.nl

Word count: 3308
Summary

For the measles-mumps-rubella (MMR) vaccine, the WHO-recommended coverage for herd protection is 95% for measles and 80% for rubella and mumps. However, a national vaccine coverage does not reflect social clustering of unvaccinated children, e.g. in schools of Orthodox Protestant or Anthroposophic identity in The Netherlands. To fully characterise this clustering, we estimated one-dose MMR vaccination coverages at all schools in the Netherlands. By combining postcode catchment areas of schools and school feeder data, each child in the Netherlands was characterised by residential postcode, primary and secondary school (referred to as school career). Postcode-level vaccination data were used to estimate vaccination coverages per school career. These were translated to coverages per school, stratified by school identity. Most schools had vaccine coverages over 99%, but major exceptions were Orthodox Protestant schools (63% in primary and 58% in secondary schools) and Anthroposophic schools (67% and 78%). School-level vaccine coverage estimates reveal strong clustering of unvaccinated children. The school feeder data reveal strongly connected Orthodox Protestant and Anthroposophic communities, but separated from one another. This suggests that even at a national one-dose MMR coverage of 97.5%, thousands of children per cohort are not protected by herd immunity.

https://doi.org/10.1017/S0950268822001455 Published online by Cambridge University Press
Introduction

The impact of childhood vaccination programmes on protecting children from infection can be measured in various ways. Common indicators are the remaining disease incidence and disease burden, and the national vaccination coverage [1]. For measles, WHO recommends a threshold value for the vaccination coverage of 95% for two doses [2], and for rubella and mumps of 80% [3]. Many EU countries have an explicit target coverage for the measles-mumps-rubella (MMR) vaccine of 95% [1]. However, outbreaks occur among unvaccinated children when they are clustered, even at high values for an average national vaccination coverage [4]. The Netherlands, in particular, have an average national vaccination coverage for MMR that is among the highest in Europe (http://data.euro.who.int/cisid/), but the substantial socio-geographic clustering of unvaccinated children results in repeated outbreaks of measles, mumps and rubella [5-7].

Clustering of unvaccinated children is often described geographically [8-13], but most contacts between children are made at school [14]. Indeed, daycare centres or schools play a major role during outbreaks in well-immunised countries [15-18]. In the Netherlands, the clustering of unvaccinated children at school is enhanced by a school system where parents are free to start and choose schools based on their own philosophy or conviction, resulting in a range of registered school identities. There are two social groups with their own school identities that were known for a lower vaccination coverage before this study: Orthodox Protestants, who live spatially clustered in the so-called Bible Belt (Supplementary Figure 1) and have a community vaccination coverage of about 60% [19, 20]; and Anthroposophics, who live spatially more dispersed and have school vaccination coverages ranging from 60% to 90% [21]. Anthroposophics are vaccine-hesitant because they find it important that children encounter and recover from childhood diseases which they perceive as not severe [22]. There is also a more diffuse group of people who fear side-effects and perceive the diseases as mild [23], but there is not one school identity linked to that group.
Vaccination coverages by school are routinely collected in various countries, including Germany and the United States [17, 18, 24-26], but reporting is frequently done by averaging school coverages locally, by county or district [25, 27]. Whereas this is very helpful for signaling which regions are at risk of infection, an average or even a local average does not fully capture the variability between schools, which is essential to fully understand the risk of outbreaks [26]. For that, we require the complete frequency distribution of school vaccination coverages.

The aim of our study was to estimate the MMR vaccination coverages for all primary and secondary schools in the Netherlands, and assess the variation in vaccination coverage in primary and secondary schools. We focussed on a single-dose rather than the WHO-recommended two-dose vaccination, because the second dose in the Netherlands is given only at nine years of age so that the one-dose coverage is a better indication of protection in primary schools. Also, a single-dose coverage better reflects willingness versus refusal to vaccinate, and shows where the completely unvaccinated children are. We discuss the implications of the observed variation for the risk of outbreaks in a population with high mean vaccination coverage.

Methods

Briefly, we assumed a steady state population of children characterised by residential postcode, primary school, and secondary school (school career). This population is constructed by use of school-level catchment data of the postcode areas in which school children live, and school feeder data of children moving from primary to secondary school. The steady-state assumption allowed us to use postcode-level coverage data for a single cohort to estimate coverages in both primary and secondary schools. We used a statistical hierarchical model, in which the vaccination status of a child
is explained by the school career, with a fixed effect of the school identities (identity career) and random effects of the two schools. The school feeder data create dependencies of coverage estimates between primary and secondary schools sharing many children, which is especially relevant if those schools have different identities.

Datasets

We used five datasets to estimate school vaccination coverages (dark green in Figure 1), and two datasets for validation of estimated coverages (Utrecht and Anthroposophic school data).

For estimation, we combined postcode-level MMR vaccination coverage data with postcode-level data of the schools that children attend (primary school and secondary school catchment data), school-level data on numbers of children moving from primary to secondary schools at (mostly) age 12 (school feeder data), and school identity data. The coverage data came from the Dutch vaccination registry Praeventis [28], of which we used the single MMR vaccination status in 2013 of 197,382 children of birth cohort 2003 at age 10. In the Netherlands, MMR is given at 14 months and 9 years of age, so focusing on a single-dose vaccination better reflects vaccine coverage across all ages in schools. The school catchment, feeder, and identity data came from the Education Executive Agency of the Ministry of Education [29], of which we used primary school catchment data from October 2012, secondary school catchment data from October 2013, and feeder data from September 2013.

Schools were categorised into 15 groups based on their identity [29]. Most identities are related to religion, but exceptions are Municipal (“Openbaar”) which are run by the local government, and General (“Algemeen bijzonder”) which is a heterogeneous group of schools with various non-religious backgrounds (e.g. Montessori, Dalton, Jenaplan). Special types of identity were “Collaborations”, which consisted of schools categorised as collaboration (“Samenwerking”) of two
or more identities among Protestant, Roman Catholic, General, and/or Municipal; and “Other”, which consisted of all schools categorized as Rest (“Overige”) in the dataset, or left blank.

For validation we used data collected by the municipal health service Utrecht (Utrecht data), which covers part of the Bible Belt and consists of actual coverage data for 488 schools; and data from a study in 2014 on vaccination hesitancy on Anthroposophic schools (Anthroposophic school data), consisting of coverage data for 11 schools [21].

Statistical analysis

Figure 1 shows how the data are used in a model to obtain the MMR coverage per school (more detail in Supplementary Material). First, the school catchment data and feeder data were combined to build the population structure, in which each child is characterized by a postcode and a school career, i.e. a primary and secondary school, one of which the child actually attends and the other which it has attended or will attend. Then, this population structure is used with the vaccination data and school identities in a Bayesian regression analysis in stan (mc-stan.org), called from R statistical software [30] with the Rstan package (mc-stan.org/rstan). Conceptually, through the regression equation (Figure 1) the observed coverages in postcode areas are explained by the school careers of the children living in that postcode area, with a fixed effect due to the identities of the schools, and a random effect for each school.

The regression analysis resulted in estimated vaccination coverages per school career, which were used to derive our results of interest: vaccination coverages per school. The analysis also provided estimated coverages per postcode area.

Because there were 15 different school identities in the data, and therefore many more possible identity careers, we had to aggregate identities and identity careers to limit the number of variables
in the regression. This was done in two preliminary analyses. The first was to define a “Rest” identity by grouping all identities without evidence of lower coverage than the major identities (Municipal, Protestant, and Roman Catholic). That resulted in five primary school identities and three secondary school identities. The second was to aggregate identity careers where possible, resulting in a final set of eight identity careers in the final analysis (see Supplementary Material for more detail).

Results

Aggregating school identities and identity careers

We first ran two preliminary regression analyses (detail in Supplementary Material) in order to reduce the number of school identities and identity careers in our final analysis. Most school identities were aggregated into one “Rest” group, but identities that remained separate were Anthroposophic and Orthodox Protestant in both primary and secondary schools, and General and Hindu in primary schools (Table 1). From these remaining identities, 15 identity careers were formed, of which eight were kept separate after a second aggregation step to obtain our final results (Table 2).

Estimating school vaccination coverages

The overall vaccination coverage in our dataset (one dose of MMR vaccine in birth cohort 2003 at the age of 10) was 97.5%. Figure 2 shows the comparison of estimated school vaccination coverages with the validation datasets. In all Utrecht schools as well as the Anthroposophic schools there is a clear correlation between true and estimated coverage, and almost all credible intervals include the
true coverage. Two Rest schools with coverages below 90% have too high estimates; these are both Protestant schools.

There is clear spatial clustering of unvaccinated children: most postcode areas had vaccination coverages around 99%, but about 4% of postcode areas had coverages below 95%, mainly in the Bible belt (Figures 3, 4). At school level, nonetheless, clustering was much more pronounced. The large majority of schools have have higher coverages than postcode area, above 99%, whereas specific schools have much lower coverages. In primary schools, vaccination coverage on Orthodox Protestant schools (62.8%) is slightly lower than on Anthroposophic schools (66.6%), whereas in secondary schools this difference is much larger (58.4% and 78.4%, respectively, Table 3). General primary schools have a slightly lower vaccination coverage than Rest primary schools (Table 3), but less than 1% of all these schools have a coverage below 95% (Figure 3). Coverages of Hindu schools are estimated to be lower than in the Rest or General schools, but this result is highly uncertain because of the small number of schools (Table 3). Of all identities, Orthodox Protestant schools have the largest variation in vaccine coverage (Figure 3).

Figure 4 shows maps of the school locations, partitioned by school identity, colour coded to indicate vaccination coverage, and as a reference a postcode map of vaccination coverages. The Orthodox Protestant schools are clearly concentrated in the Bible Belt, with most of the very-low coverage schools east from the centre. The lower-coverage Rest schools are also predominantly located in the Bible Belt area; most of these have the Protestant identity (not shown). Anthroposophic schools are more widely distributed across the country.

Clustering of unvaccinated children
In our age cohort dataset of 197,382 10-year old children in the Netherlands in 2013, 4998 (2.5%) were not vaccinated against MMR. In primary schools, 51% of unvaccinated children attend General or Rest schools most of which have a very high vaccination coverage, but the rest is clustered in schools with much lower coverages: 38% in Orthodox Protestant schools, 10% in Anthroposophic schools, and 1% in Hindu schools (Figure 3, Table 3). When these children go to secondary school, this hardly changes: 56% go to Rest, 39% to Orthodox Protestant, and 5% to Anthroposophic schools. Many children move to secondary schools of the same identity as their primary school, more so in the Orthodox Protestant community (85%) than in the Anthroposophic community (44%), showing that these children go to low-coverage schools during their whole school career. There were, however, no children moving between schools with these two identities.

Discussion

Our aim was to estimate single-dose MMR vaccination coverages on schools in the Netherlands, and assess clustering of unvaccinated children in schools. It turns out that about half of the unvaccinated children are clustered into a small minority of schools with very low vaccination coverage. These are almost all schools with an Orthodox Protestant or Anthroposophic identity. That implies that thousands of children per cohort are not protected by vaccination, even when the mean vaccination coverage of 97.5% is well above WHO recommendations for both infections.

Spatial clustering of unvaccinated children in the Bible belt is clearly observed at the postcode area level (Figure 4), and has been described before. Our results, however, show that clustering is much stronger at the level of schools, which better reflects the social environment of children. Most schools have vaccination coverages higher than in most postcode areas, whereas some schools of mainly Orthodox Protestant and Anthroposophic identity have much lower coverages (Figure 3). The analysis also revealed social clustering outside the Bible belt area (Figure 4). Apart from the good
agreement with the validation data, the Orthodox Protestant coverages also confirm estimates of 60% in this community from online surveys by Ruijs et al [19].

Whereas Orthodox Protestant and Anthroposophical schools are only minority identities, about half of all unvaccinated children attend these schools with low vaccination coverage (Table 3). Children not only spend much of their time at school, their social network outside school is probably very similar. That creates an infection risk for many unvaccinated children, who seem to be protected when considering their residential four-digit postcode area (Figure 5), let alone the national vaccination coverage. However, for virus introductions to cause large epidemics, there should also be sufficient contacts between the schools. Most of these contacts are probably through households, mainly between primary and secondary schools. Aside from the geographic proximity (Figure 2), the school identity careers (Table 2) show that the Orthodox Protestant community is closely knit: 85% of children attending Orthodox Protestant primary schools, also attend secondary school with that identity, and vice versa. In addition, families in the Orthodox Protestant community tend to be large [31], creating more links between primary and secondary schools. This may explain the occasional large epidemics mostly within this community. Because secondary schools are essential links in the school network, large epidemics may only occur when the interepidemic period is long enough so that these schools are populated with susceptible children. This is consistent with the pattern of measles epidemics in the Bible Belt in the Netherlands, which took place in 1988, 1999, and 2013 [7, 32].

Anthroposophic schools are less well connected, firstly because they are more evenly distributed geographically (Figure 2), and secondly because more children going to either an Anthroposophic primary or an Anthroposophic secondary school but not both (Table 2). Indeed, outbreaks in Anthroposophic schools tend to stay confined to a few schools [33]. Also, the school feeder data suggest that the two low-coverage communities are socially separated, as there were no children in
the dataset moving between Orthodox Protestant and Anthroposophic schools (Table 2). As a consequence, during the epidemics in the Orthodox Protestant community, Anthroposophic schools were hardly affected [7].

The estimated school vaccination coverages captured the trend of the validation datasets well, but were too unreliable to report for individual schools (Figure 2). First, they were unprecise, which is reflected by the wide credible intervals, and caused by the large number of schools compared to postcode area coverage data. Second, in a few cases estimates were inaccurate, notably for two schools of the Rest identity (Figure 2), which were Protestant by their registered identity (before aggregation of identities). Misclassification of some Protestant school is likely: the websites of these schools reveal an Orthodox identity, and most low-coverage Rest schools in the Bible Belt (Figure 4) are also Protestant. Third, confounding may have resulted in the estimated low coverages of the Hindu schools, because these schools are exclusively situated in large cities with neighbourhoods with lower socio-economic status and a more diverse ethnic background which are associated with lower vaccine uptake [34]. The statistical procedure might have created an incorrect association between lower coverage and these schools, as a lower coverage was not associated with the Hindu religious group in a large cross-sectional survey [34], in which 32 out of 35 Hindu respondents of age 1.5-20 were vaccinated (Mollemma, personal communication). Fourth, we used data of a single birth cohort, which may not be representative if the vaccination coverage has increased or decreased, and which may overestimate coverage in younger children if they receive their first MMR vaccination at the age of nine, when the routine second vaccination is offered. If reliable coverages could be obtained, e.g. by coupling of school and vaccination registries, they could be used to inform parents, to prioritize supplemental vaccinations during outbreaks, or for more detailed studies into changes in vaccine hesitancy in different social groups, represented by school identity.
Pockets of low vaccination coverage and resulting risk of outbreaks are better identified in social than geographic environments [35]. We assessed heterogeneity in vaccine coverage by estimating coverages at schools. The reason is that schools are good reflections of the social environment of children, especially if parents can choose a school with a clear identity like in the Netherlands. For the Netherlands, we were able to estimate school-level vaccination coverages by relating available postcode-level coverages to schools through the school catchment data, and we could use school identity to explain much of the variation between schools. In combination with the school feeder data, we could distinguish two well-connected by mutually separated clusters of schools with low vaccination coverage. For other countries, the number of postcode areas may be larger than the number of schools, which would simplify analyses, or alternative variables such as public and private schools may be more useful. Ideally, school vaccination coverage data are directly available: for instance, in the United States these are available at school level for some states [24], but often grouped by county for analysis [27]; also in Germany they are collected at school entry and then grouped by district [25]. For describing connections between schools, feeder data may be supplemented (or replaced) by individual level household data to build larger networks of schools and identify communities at risk.

Acknowledgements: We would like to acknowledge Frits Woonink (GGD Utrecht) and Wim Noort (GGD Noord-Oost Gelderland) for providing and disclosing the Utrecht and Anthroposophic school coverage data, Marc Meurs (DUO) for help with the DUO school data, and James Munday (London School of Hygiene and Tropical Medicine), Liesbeth Mollema (RIVM), Alies van Lier (RIVM), and Hester de Melker (RIVM) for discussions on data and interpretation.

Financial support: This research received no specific grant from any funding agency, commercial or not-for-profit sectors.
Conflicts of interest: None

Data availability statement: the data that support the findings of this study can be requested from the corresponding author, DK. The data are not publicly available because of privacy reasons (low number entries and individual school data).

References


https://doi.org/10.1017/S0950268822001455 Published online by Cambridge University Press


Gromis A, Liu KY. Spatial Clustering of Vaccine Exemptions on the Risk of a Measles Outbreak. *Pediatrics* 2022; 149(1).


Table 1. Numbers of primary and secondary schools by identity

<table>
<thead>
<tr>
<th>Identity</th>
<th>Primary schools</th>
<th>Secondary schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Aggregated identity(^a)</td>
</tr>
<tr>
<td>Municipal</td>
<td>2444</td>
<td>Rest</td>
</tr>
<tr>
<td>General</td>
<td>497</td>
<td>General</td>
</tr>
<tr>
<td>Anthroposophic</td>
<td>70</td>
<td>Anthroposophic</td>
</tr>
<tr>
<td>Roman Catholic</td>
<td>2236</td>
<td>Rest</td>
</tr>
<tr>
<td>Protestant</td>
<td>1882</td>
<td>Rest</td>
</tr>
<tr>
<td>Reformed (Liberated)</td>
<td>118</td>
<td>Rest</td>
</tr>
<tr>
<td>Orthodox Protestant</td>
<td>180</td>
<td>Orthodox Protestant</td>
</tr>
<tr>
<td>Evangelic</td>
<td>13</td>
<td>Rest</td>
</tr>
<tr>
<td>Evangelical Fraternity</td>
<td>2</td>
<td>Rest</td>
</tr>
<tr>
<td>Hindu</td>
<td>6</td>
<td>Hindu</td>
</tr>
<tr>
<td>Islamic</td>
<td>43</td>
<td>Rest</td>
</tr>
<tr>
<td>Jewish</td>
<td>2</td>
<td>Rest</td>
</tr>
<tr>
<td>Interconfessional</td>
<td>12</td>
<td>Rest</td>
</tr>
<tr>
<td>Collaborations</td>
<td>65</td>
<td>Rest</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>7570</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) identities after aggregation: identities with vaccination coverage not different from Municipal, Roman Catholic, and Protestant were aggregated into Rest
Table 2. Numbers of children in feeder data, by school identity career.

<table>
<thead>
<tr>
<th>Primary school identity</th>
<th>Secondary school identity</th>
<th>Anthro.</th>
<th>Orthodox</th>
<th>Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ant - Ant</td>
<td>0</td>
<td>-</td>
<td>876</td>
</tr>
<tr>
<td></td>
<td>Ort - Ort</td>
<td>4298</td>
<td>765</td>
<td></td>
</tr>
<tr>
<td>Hindu</td>
<td>0</td>
<td>0</td>
<td>262</td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>Gen - Rest</td>
<td>3</td>
<td>13246</td>
<td></td>
</tr>
<tr>
<td>Rest</td>
<td>Rest - Rest</td>
<td>782</td>
<td>182471</td>
<td></td>
</tr>
</tbody>
</table>

*identity careers after aggregation: identity careers with vaccination coverage not different from Rest-Rest were merged with a similar identity career. Abbreviations: Ant = Anthroposophic, Gen = General, Hin = Hindu, Ort = Orthodox Protestant*
Table 3. Vaccination status and number of unvaccinated children per age cohort, by school identity.

<table>
<thead>
<tr>
<th>School identity</th>
<th>Primary schools</th>
<th></th>
<th>Secondary schools</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean vaccination coverage per school(^a)</td>
<td>Unvaccinated children per cohort(^b)</td>
<td>Mean vaccination coverage per school(^a)</td>
<td>Unvaccinated children per cohort(^b)</td>
<td></td>
</tr>
<tr>
<td>Anthroposophic</td>
<td>66.6% [59.6 ; 73.4]</td>
<td>496 [402 ; 588]</td>
<td>78.4% [71.3 ; 85.5]</td>
<td>242 [164 ; 321]</td>
<td></td>
</tr>
<tr>
<td>Orthodox</td>
<td>62.8% [60.3 ; 65.2]</td>
<td>1904 [1799 ; 2009]</td>
<td>58.4% [52.5 ; 64.0]</td>
<td>1958 [1853 ; 2063]</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>74.5% [50.3 ; 94.6]</td>
<td>60 [15 ; 95]</td>
<td>58.4% [52.5 ; 64.0]</td>
<td>1958 [1853 ; 2063]</td>
<td></td>
</tr>
<tr>
<td>Hindu</td>
<td>97.4% [96.4 ; 98.1]</td>
<td>340 [243 ; 448]</td>
<td>98.4% [98.3 ; 98.5]</td>
<td>340 [243 ; 448]</td>
<td></td>
</tr>
<tr>
<td>Rest</td>
<td>98.7% [98.6 ; 98.8]</td>
<td>2198 [2041 ; 2366]</td>
<td>98.4% [98.3 ; 98.5]</td>
<td>2797 [2655 ; 2942]</td>
<td></td>
</tr>
<tr>
<td>All schools</td>
<td>97.5% [97.4 ; 97.5]</td>
<td>4998 [4872 ; 5129]</td>
<td>97.5% [97.3 ; 97.7]</td>
<td>4998 [4872 ; 5129]</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) posterior mean coverage in all children in the group of schools, and 95% credible interval

\(^b\) posterior estimated number of unvaccinated children in each age cohort, and 95% credible interval
Figure 1. Conceptual outline of the method to estimate school vaccination coverages (full details in Supplementary Material). Numbers are fictional but mutually consistent, e.g. population structure $C$ contains all children in datasets $P$ and $S$ (equations (1), (2), (3) in Supplementary Material). Green tables are data: dark green are original datasets, light green are linked data. Blue tables are results: dark blue are original estimates (logit-transformed in the actual analysis), light blue are derived quantities. Top (data): three population datasets are linked to build the population structure of numbers of children by school career and residential postcode area.
Middle (model): regression model with vaccination data per postcode as dependent variable; the independent variables are the school careers of the children, mapped to postcode areas through function f(.), in which the population structure is used to map school career coverages to postcode area coverages, and in which the logit(coverage) of each school career is modelled as the sum of a fixed identity career effect and two random school effects. Bottom (results): the estimated coverages and deviations are used to calculate the coverage per school career, and combined with the population structure to calculate the coverages per school and per postcode.

Figure 2: validation plots of estimated vaccination coverages with 95% credible interval compared to actual vaccine coverage in Utrecht, 2013 (left) and on Anthroposophical schools in the Netherlands, 2012 (right).
Figure 3: posterior median vaccination coverages of all schools, grouped by school identity.

Displayed are medians (of the median coverages) with interquartile range of (IQR, boxes), and whiskers extending to 1.5 times the IQR or most extreme value, whichever is closest to the median. Dots are outliers.
Figure 4: geographic distributions of all schools by identity, with colour coded vaccination coverage. The bottom right map shows the MMR coverages by four-digit postcode area.