Universiteit
Leiden
The Netherlands

A solid start for the Dutch first thousand days-approach: insights into program adoption, monitoring and cross-sectoral collaboration

Molenaar, J.M.

Citation
Molenaar, J. M. (2024, September 4). *A solid start for the Dutch first thousand days-approach: insights into program adoption, monitoring and cross-sectoral collaboration*. Retrieved from https://hdl.handle.net/1887/4054887

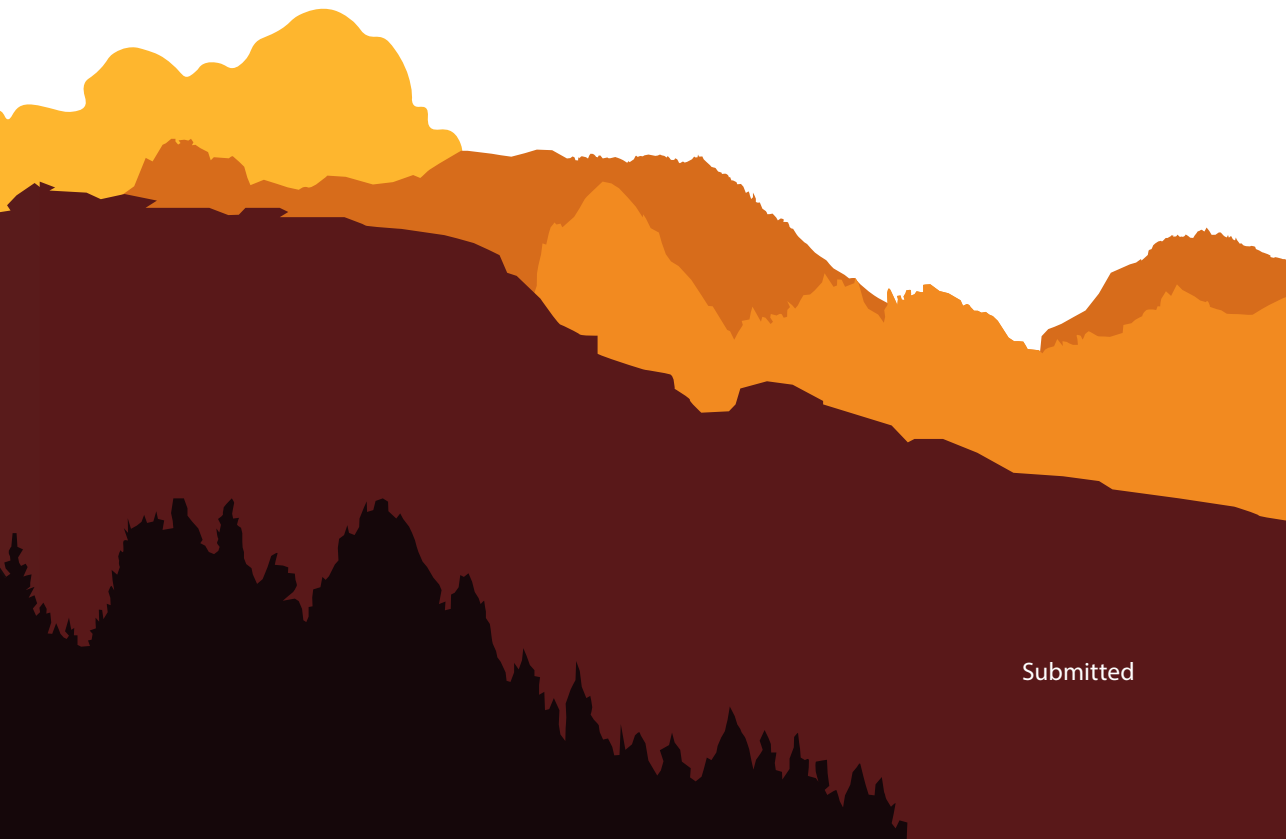
Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/4054887

Note: To cite this publication please use the final published version (if applicable).

3

Predicting population-level vulnerability among pregnant women using routinely collected data and the added relevance of self-reported data

Joyce M. Molenaar
Ka Yin Leung
Lindsey van der Meer
Peter Paul F. Klein
Jeroen N. Struijs
Jessica C. Kiefte-de Jong



Submitted

ABSTRACT

Background
Recognizing and addressing vulnerability during the first thousand days of life can prevent health inequities. It is necessary to determine the best data for predicting multidimensional vulnerability (i.e. risk factors to vulnerability across different domains and a lack of protective factors) at population-level to understand national prevalence and trends. This study aimed to 1) assess the feasibility of predicting multidimensional vulnerability during pregnancy using routinely collected data, 2) explore potential improvement of these predictions by adding self-reported data on health, wellbeing and lifestyle, and 3) identify the most relevant predictors.

Methods
The study was conducted using Dutch nationwide routinely collected data and self-reported Public Health Monitor data. First, to predict multidimensional vulnerability using routinely collected data, we used Random Forest (RF) and considered the Area Under the Curve (AUC) and F1-measure to assess RF-model performance. To validate results, sensitivity analyses (XGBoost and Lasso) were done. Second, we gradually added self-reported data to predictions. Third, we explored the RF-model's variable importance.

Results
The initial RF-model could distinguish between those with and without multidimensional vulnerability (AUC 0.98). The model was able to correctly predict multidimensional vulnerability in most cases, but there was also misclassification (F1-measure 0.70). Adding self-reported data improved RF-model performance (e.g. F1-measure 0.80 after adding perceived health). The strongest predictors concerned self-reported health, socioeconomic characteristics and healthcare expenditures and utilization.

Conclusions
It seems possible to predict multidimensional vulnerability using routinely collected data that is readily available. However, adding self-reported data can improve predictions.

INTRODUCTION

A strong foundation during the first thousand days of life, which span from conception till a child's second birthday, can positively impact health and development in later life and across generations (1, 2). Adverse experiences and exposures can influence the health of parents themselves, but can also be transmitted to their children, which, as these children grown into adulthood and potentially become parents themselves, leads to new cycles of adversity (2). In order to prevent health inequities and break the intergenerational cycle, it is important to recognize and address vulnerability during the first thousand days (1-3). This is also a focus in the Dutch nationwide action program Solid Start (in Dutch: Kansrijke Start) (4). The concept of vulnerability is often used to describe subgroups with increased risks to adverse health outcomes or limited access to healthcare. In short, vulnerability encompasses a multifaceted and dynamic process in which diverse stressors at the individual or contextual level can serve as risk factors, whereas protective factors have the potential to mitigate or prevent vulnerability (5-9). Examples of risk factors encompass unemployment or stress, while examples of protective factors include a strong social network or effective coping skills.

The concept of vulnerability and its scope has garnered increasing attention among providers and policymakers who strive to enhance the provision of care and support during the first thousand days of life (4, 9, 10). In daily care, a common understanding between professionals from the medical and social sector on the characteristics of high-risk individuals can foster mutual understanding and improve cross-sectoral collaboration (9). At national and local policy levels, drawing attention to the prevalence, geographical distributions and trends in vulnerability can support policy monitoring and prioritization. These insights not only foster a sense of urgency, but also enhance the conversation between different stakeholders, and facilitate vision formulation and intervention prioritization (11).

This study extends our prior research to predict population-level vulnerability among pregnant women. Our previous study highlighted the significance of considering both risk and protective factors, particularly in the context of adverse outcomes (12). Through Latent Class Analysis (a data-driven technique to identify subgroups with similar characteristics), we identified five groups of pregnant women with different social risk and protective factors to vulnerability prior to pregnancy. Women in the 'multidimensional vulnerability'-group shared multiple risk factors across several domains (e.g. psychosocial, medical, and socioeconomic), lacked protective factors and were most at risk of adverse outcomes such as premature birth and caesarean section. Having risk factors in a single domain (e.g. socioeconomic) was not necessarily associated with adverse outcomes. This study utilized both routinely collected observational data and self-reported data on health, wellbeing and lifestyle of the Public Health Monitor 2016 (PHM-2016) to predict multidimensional vulnerability (12). Using the PHM-2016 resulted in a subset of the total Dutch pregnant population. Hence, the prevalence of multidimensional vulnerability across the entire population of pregnant women at a national level remains unknown, and it is unclear

whether this can be assessed using solely routinely collected observational data and what the added value of self-reported data is. Moreover, we lack an understanding of the strongest predictors for population-level vulnerability.

Mapping out the percentage of multidimensional vulnerability among pregnant women in the Netherlands and its predictors is relevant for risk stratification. In population health management, this is an essential initial step to tailor (preventive) actions to the needs of specific risk-groups to enhance population health (13, 14). Such stratification commonly relies on routinely collected data (15), offering advantages such as widespread availability, reduced practitioner burden, time and costs (16, 17). Moreover, the longitudinal and systematic approach facilitates comparisons over time (16, 17). However, it is important to empirically evaluate whether routinely collected data is sufficient for risk-stratification for high-risk groups. In addition, there is a potential for improvement in predicting multidimensional vulnerability at population-level by incorporating self-reported health, well-being and lifestyle data. For example, studies indicate that self-reported health and vulnerability correspond to or complement clinical measures in predicting adverse health outcomes (18-22). Yet, the impact of adding self-reported data next to routinely collected data in predicting vulnerability remains unexplored.

This study has three objectives. First, to assess the feasibility of accurately predicting multidimensional vulnerability during pregnancy at population-level using solely routinely collected observational data. We use the predictions to report on the prevalence and spatial variation of multidimensional vulnerability during pregnancy at population-level in the Netherlands. Second, to identify whether self-reported data on health, wellbeing, and lifestyle could improve those predictions with routinely collected data. Third, to identify the predictors that have the most significant impact on the classification of multidimensional vulnerability.

METHODS

Data sources

This study employed data from DIAPER (Data-InfrAstructure for ParEnts and childRen) (17). DIAPER integrates individual-level, routinely collected observational data from various nationwide data sources in the Netherlands, including Perined and Statistics Netherlands. Perined collects routine care data on pregnancy, birth, and the first 28 days after birth from midwives, gynaecologists, and paediatricians (23). Statistics Netherlands collects data about social issues, including health, welfare, income, education, and labour (24, 25). To enrich DIAPER, self-reported data on health, well-being, and lifestyle of the PHM-2016 were included (26). The PHM is a health survey conducted every 4 years among a varying sample of Dutch adults aged 19 years and older (about 450.000 in 2016).

Study population

The study population consisted of 4172 unique women with a pregnancy and childbirth in 2017 or 2018 who participated in the PHM-2016 prior to pregnancy. Details on selecting the study population are described in Chapter 2 (12). To illustrate the prevalence and spatial variation of multidimensional vulnerability at national level, all unique registered pregnancies in Perined from 2017 to 2021 were considered (n = 807.904) (17). Missing data were imputed through Multiple Imputation using Chained Equations (MICE), leading to five imputed datasets (27).

Independent variables

Analogous to our previous study, we included 42 variables in the predictive models (12). Details on these variables have been described in Appendix 1 of our previous study (Chapter 2) (12). The first category in each variable denotes the risk factor to vulnerability.

Of those 42 variables, 31 variables concerned routinely collected data available for all pregnant women in DIAPER (n = 807.904). Those included individual characteristics (age, ethnicity, parity, asylum seeker status), socioeconomic characteristics (educational level, household income, socioeconomic position by occupational status, debts and payment arrears, permanent employment contract, and full-time employment contract), household characteristics (type of household, marital status, dissolution of marriage, household size, and youth support utilization), healthcare expenditures and utilization (total healthcare expenditures, General Practitioner's (GP) expenditures, hospital expenditures, medication use, and addiction related care utilization), psychosocial characteristics (mental healthcare utilization, mild intellectual disability), life events (crime suspect, crime victim, having been detained, frequent moving, loss of a family member), living conditions (home ownership, motorized vehicle ownership, proximity to GP office, liveability neighbourhood).

The other 11 variables were derived from the PHM-2016 and consequently only available for 4172 individuals. These variables included lifestyle factors (smoking, alcohol use, physical activity, Body Mass Index (BMI)), self-reported health (perceived health status, long-term illness, restricted by health), psychosocial characteristics (risk of depression or anxiety disorders, loneliness, feelings of control over life) and socioeconomic characteristics (insufficient financial resources).

Outcome: multidimensional vulnerability

The outcome measure is multidimensional vulnerability, as derived from our previous study (Chapter 2) (12). Women classified into the 'multidimensional vulnerability'-class share a combination of multiple risk factors to vulnerability in several domains and lack protective factors. It is not a straightforward equation and risk factors vary across individuals. Most present risk factors include not having an income or receiving benefits, rental housing, high GP healthcare expenditures, long-term illness, negative self-perceived health, and elevated risks of feeling lonely, depressed or anxious.

We added the variable multidimensional vulnerability to the dataset of 4172 individuals. All women who were previously assigned to the multidimensional vulnerability-class were classified as 'yes' (n = 249) and women in all other classes as 'no' (n = 3923).

Statistical analyses

To assess whether it is feasible to predict multidimensional vulnerability during pregnancy using solely routinely collected data at population-level (objective 1), we employed Random Forest (RF). RF is a machine learning method for regression and classification that operates through the construction of multiple decision trees (28). The method makes no assumptions about data distribution and works well with the number of individuals in our dataset relative to the number of variables. Sensitivity analyses were conducted using XGBoost and Lasso for validation (see Appendix 1).

We sought for the optimal model using the Area Under the Curve (AUC) and F1-measure (29). The AUC, ranging between 0.5 (random) to 1.0 (perfect model), illustrates the ability of the model to distinguish between those with and without multidimensional vulnerability. Due to our imbalanced dataset with relatively few cases of multidimensional vulnerability, we calculated F1-measures to focus on correct predictions of vulnerability (29). The F1-measure balances precision, also known as positive predictive value (i.e. proportion of correct predictions out of all predicted as vulnerable) and recall/sensitivity (i.e. proportion of individuals with vulnerability correctly predicted as vulnerable by the model). We treated both elements as equally important. A perfect score means the model can identify *all* positive cases while also identifying *only* positive cases (instead of assigning those without vulnerability incorrectly to the vulnerability-class). We additionally report on specificity (i.e. proportion of correct negative predictions out of all without vulnerability) and the confusion matrices showing true/false positives and true/false negatives. In model development, we used default hyperparameters settings in the R-packing 'ranger' (30), as these typically perform well. We used nested cross-validation to choose the threshold probability for classifying multidimensional vulnerability into 'yes' and 'no' and to assess model performance (31). This involved splitting the dataset in an outer loop (six folds of train-test combinations) and inner loop (five train-validate combinations), detailed in Appendix 1. The final RF-model can be utilized for predicting outcomes on new datasets. Being the best performing model, it was also used to report on the prevalence and spatial variation of multidimensional vulnerability during pregnancy from 2017 to 2021. We computed percentages for both national and municipality levels in the five imputed datasets and we conducted an additional complete cases analysis at national level for comparison. Municipality level results were visualized on a map of the Netherlands.

Next, to identify if self-reported data on health, wellbeing and lifestyle could improve predictions with solely routinely collected data (objective 2), we gradually added self-reported data from the PHM-2016 to the RF-model. Using the previous six train-test combinations, we calculated average F1-measures for different variable sets; 1) solely routinely collected data (baseline, 31 variables); 2) baseline combined with one varying PHM-2016 variable (comprising 32 variables); 3) baseline combined with two varying

PHM-2016 variables (comprising 33 variables); and 4) baseline combined with all PHM-2016 variables, representing a potential optimum (42 variables). Comparing average F1-measures for each combination helped identify which PHM-2016 variables enhanced model performance.

To identify which variables were most important in model predictions (objective 3), we assessed variable importance in the final RF-model with and without PHM-2016 data. Variable importance was measured using out-of-bag (OOB) observations, explained in Appendix 1. This process yields a ranking of variable importance (32). As sensitivity analyses, we checked the permutation importance and Partial Dependence Plots (PDPs), explained in Appendix 1.

Ethics approval

The Clinical Expertise Centre of the National Institute for Public Health and the Environment confirmed that our study was not subject to the Dutch Medical Research involving Human Subjects Act (WMO) (reference number: VPZ-574).

RESULTS

Study population

The study population comprised 4172 women (Appendix 2). Approximately 42.1% of these women were nullipara, 4.6% had a low income and 6.0% a low educational level. In comparison to all women with unique pregnancies between 2017 and 2021 (n = 807.904), the distribution regarding most variables was comparable, but differences were found for variables such as income, educational level and ethnicity. Among the 4172 women, there was generally a lower incidence of the risk factors.

Predictions with routinely collected data

The RF-model which included the routinely collected data obtained an average AUC of 0.98 (see Table 1). Such a high AUC implicates that the model sufficiently distinguishes between those with and without multidimensional vulnerability. The F1-measure had an average of 0.70, indicating that the model is able to correctly predict cases of multidimensional vulnerability, but that there are also cases missed as well as women incorrectly assigned to the vulnerability-class. Appendix 2 presents the selected hyperparameters and thresholds and the results of the separate folds. Results were consistent with those of XGBoost and Lasso (Appendix 2).

The percentage of individuals with multidimensional vulnerability during pregnancy in the Netherlands was 8.1 in 2017 and decreased to 7.2 in 2021, as derived from the RF-model (Figure 1). The percentages were slightly higher for XGBoost and lasso (respectively 8.0% and 9.1% in 2021), but showed a similar decreasing trend, as printed in Appendix 2. Appendix 2 additionally shows the complete case analysis.

Figure 2 visualizes the geographical distribution of multidimensional vulnerability during pregnancy in the Netherlands over the years 2017 to 2021, based on predictions of the RF-model. There are differences between municipalities, with percentages ranging from 1.8 to 17.5%.

Table 1. Results of the RF and the sensitivity analyses

	Metrics Mean from five-fold cross validation (SD)					**Confusion matrices for best fold** Number in each category	
	AUC	F1-measure	Precision	Recall/ sensitivity	Specificity		
Random Forest	0.98 (0.00)	0.70 (0.03)	0.74 (0.06)	0.66 (0.04)	0.98 (0.00)	30 (TP) 6 (FP)	14 (FN) 645 (TN)
XGBoost	0.98 (0.00)	0.68 (0.04)	0.70 (0.02)	0.67 (0.08)	0.98 (0.00)	34 (TP) 10 (FP)	13 (FN) 638 (TN)
Lasso regression	0.98 (0.01)	0.68 (0.04)	0.67 (0.07)	0.70 (0.07)	0.98 (0.01)	32 (TP) 12 (FP)	11 (FN) 640 (TN)

AUC = Area Under the Curve, TP = True Positive, FN = False Negative, FP = False Positive, TN: True Negative
Results based on analyses among study population of 4172 women

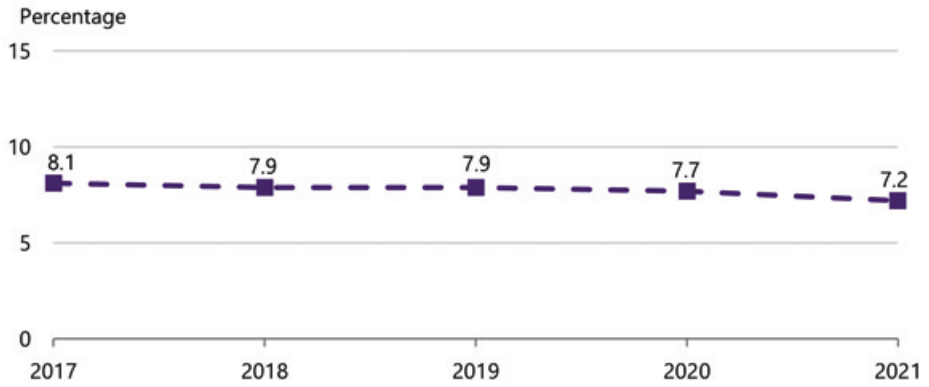


Figure 1. Percentage of multidimensional vulnerability during pregnancy in the Netherlands during the years 2017 to 2021, based on the RF-model using routinely collected data prior to pregnancy.
Results based on analyses among all unique pregnancies from 2017 – 2021 (n = 807.904)

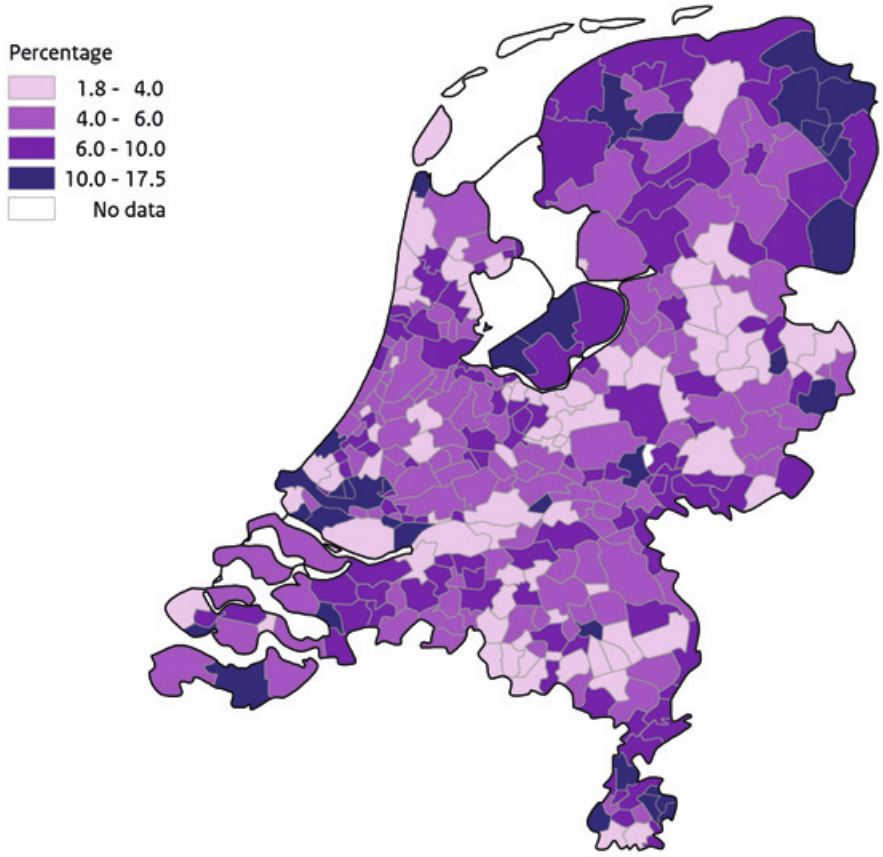


Figure 2. Heatmap visualizing the geographical distribution of multidimensional vulnerability during pregnancy in the Netherlands, at municipality level, for all pregnancies from 2017 to 2021. A darker color indicates a higher percentage of vulnerability.
Results based on analyses among all unique pregnancies from 2017 – 2021 (n = 807.904)

Adding self-reported data to predictions

The baseline F1-measure (using routinely collected data; 31 variables) was 0.70 and the potential optimum (using both routinely collected data and all self-reported data of the PHM-2016; 42 variables) was found to be 0.83, shown as vertical lines in Figure 3. Including self-reported variables improved the performance of the RF-models with solely routinely collected data. Especially self-reported data on 'perceived health status' (average 0.80) and 'restricted by health' (0.79) improved the model's performance, but also 'long-term illness' (0.77) and 'risk to depression or anxiety disorders' (0.74). Others had little impact or slightly decreased performance, such as physical activity. Appendix 2 presents the results of adding two varying self-reported variables. This further improved the performance of the model.

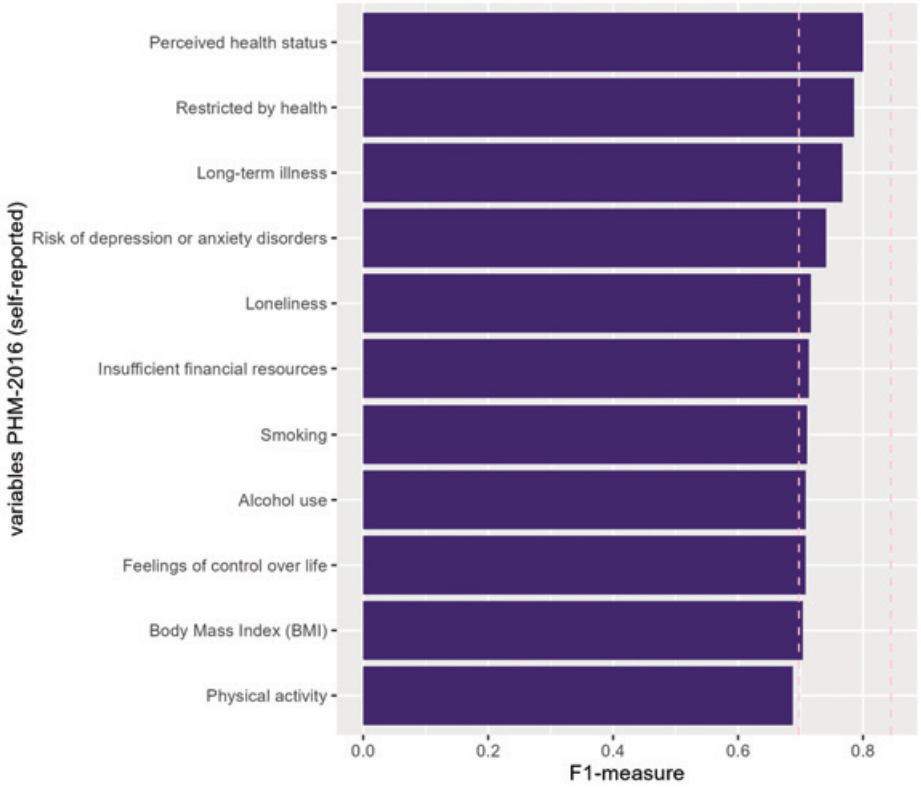


Figure 3. Variables on self-reported health, wellbeing and lifestyle added to the RF-model with solely routinely collected data. The vertical lines show the average F1-measures.
Results based on analyses among study population of 4172 women

Variable importance

Figure 4 shows the variable importance of all 42 variables. Top seven predictors for multi-dimensional vulnerability during pregnancy were: 'socioeconomic position (occupational-status)', 'perceived health status', 'restricted by health', 'permanent employment contract', 'medication use', 'long-term illness' and 'total healthcare expenditures'. Out of these seven variables, which represent both risk and protective factors, three concern self-reported health, two concern socioeconomic characteristics and two relate to healthcare expenditures and utilization. Related to financial status, self-reported 'insufficient financial resources' was ranked higher compared to the routinely collected 'household income' and 'depts and payment arrears'. Likewise, self-reported 'perceived health status' and 'feeling restricted by health' was ranked higher than 'medication use' and 'total healthcare expenditures'. We found the opposite for psychological characteristics: routinely collected 'mental healthcare utilization' was ranked higher than self-reported 'risk of depression or anxiety disorders' or 'loneliness'. However, differences were small.

The permutation importance ranking (Appendix 2) yielded comparable results, although 'mental healthcare utilization' and 'GP expenditures' were ranked slightly higher. Appendix 2 additionally shows the rankings without self-reported data, using solely routinely collected data (31 variables).

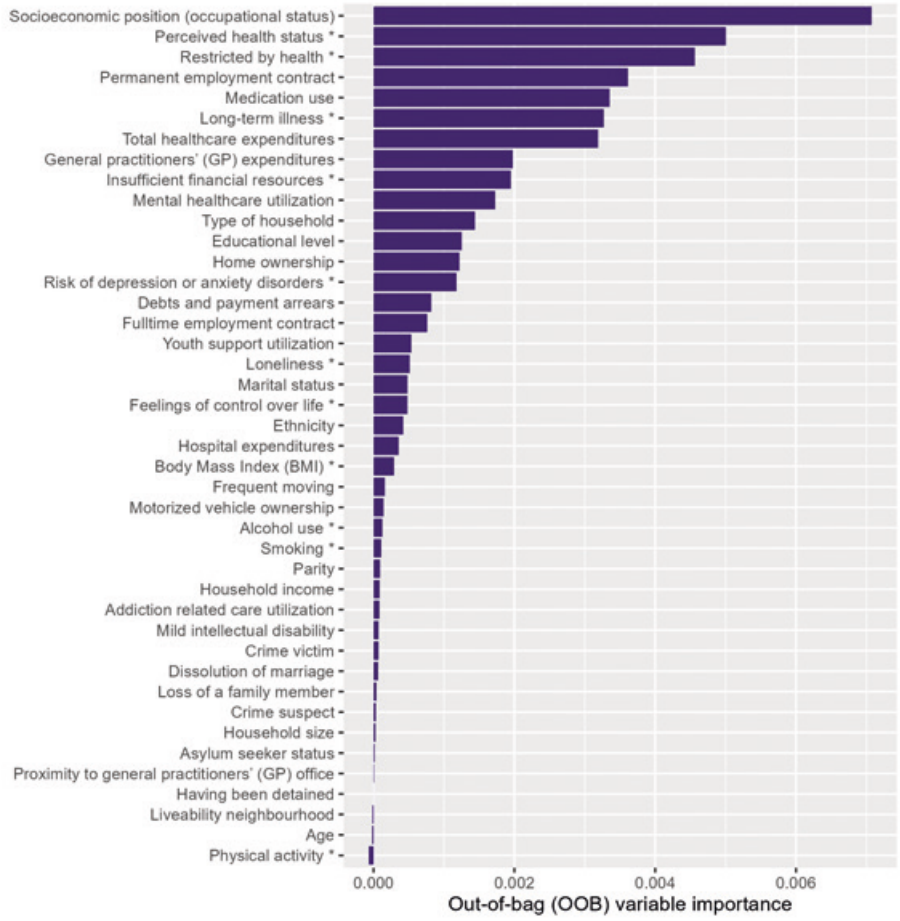


Figure 4. Variable importance ranking of the RF-model for 'multidimensional vulnerability'. The set of 42 variables used for classification are ordered from high to low importance. The length of a line represents the importance of a particular variable on the model's predictions. Self-reported data of the PHM-2016 is indicated with an asterisk (*).
Results based on analyses among study population of 4172 women

DISCUSSION

This study provides insight into predicting multidimensional vulnerability during pregnancy at population-level in the Netherlands using pre-pregnancy routinely collected data and the relevance of additional self-reported data on health, wellbeing, and lifestyle. Based on our results, it seems reasonably feasible to predict multidimensional vulnerability using solely routinely collected data, since the RF-model could distinguish between those with and without multidimensional vulnerability and was able to correctly predict multidimensional vulnerability in many cases. However, we found that adding self-reported data improved model performance. Out of the seven strongest predictors to multidimensional vulnerability in our dataset, three concerned self-reported health, two concerned socioeconomic characteristics, and two related to healthcare expenditures and utilization.

Using solely routinely collected data to predict multidimensional vulnerability appears feasible, but several women were wrongly assigned to the vulnerability class, and other cases were missed. The crucial concern is whether the model achieved adequate performance, prompting consideration of using this readily available routinely collected data versus acquiring self-reported data on experienced health. Both data sources have advantages and disadvantages, and may be used for different purposes. Using routinely collected data is relatively easy, accessible and time efficient. This pragmatic approach recognizes that not all data are available and can be utilized, analysed and interpreted. However, it is less accurate which might mainly affect those missed by the model. Considering all relevant factors by using additional self-reported data leads to better predictions. However, this has numerous implications and inherent challenges, including increased burden to practitioners, time and costs. Based on our study, we consider routinely collected data sufficient for policy monitoring of multidimensional vulnerability at population-level. It can offer insight into its scope and development over the years and help identify municipalities and neighbourhoods characterized by increased vulnerability, enabling tailored (preventive) measures for efficient budget allocation. Simultaneously, we agree with previous scholars that applying vulnerability in a dichotomous way is challenging as the concept is multi-layered, contextualized and dynamic, requiring caution to avoid over-inclusion or exclusion of individuals (33, 34). Our previous study (12) revealed a greater array of vulnerability groups, with women having risk factors within one specific domain and protective factors in others. We must not overlook these and other intermediary and personal, contextual forms of vulnerability. Our predictive RF-model was not intended for application in individual predictions and individual decision-making but meant for risk-stratification on a population-level. Because risk assessment is not straightforward, we consider routinely collected data by itself unsuitable for individual predictions, given that it insufficiently accounts for protective factors and coping strategies at an individual level, among others. We believe that an open conversation with (future) parents about their experienced health and well-being is indispensable to better understand their context and needs. It is essential that this is accompanied by a trusting relationship, and appropriate follow-up steps, preventing stigmatization, simplification

and harm (22, 33, 35). Given the added value of self-reported data however, we suggest to explore how perceived health can be systematically included into screening guidelines and care registries for professionals, to enhance the provision of personalized care and support while further improving population-level predictions in the future.

In our study, adding self-reported data led to better model performance and self-reported health indicators were found as important predictors to multidimensional vulnerability. Consistent with the psychosocial literature, several subjective measures (e.g. self-reported 'insufficient financial resources') outweighed objective measures (e.g. registered 'income' and 'debts and payment arrears') as predictors in our study. For example, multiple studies reported a stronger link between people's subjective SES and wellbeing and physical health compared to objective SES based on income or education (36-39). Arguably, perceiving your circumstances through the lens of limited resources impacts decision-making and behaviour (e.g. favouring short-term over long-term considerations), increases uncertainties and stress, and thus exacerbates pre-existing vulnerabilities (40-43). Other studies reported how self-reported health or vulnerability correspond to, outperform or complement clinical measures in predicting physical health and mortality (18-20). However, using self-reported health also has its challenges. For instance, it provides little guidance regarding what respondents consider when reporting 'poor health' and whether they refer to physical pain, mental wellbeing, less vitality or other factors (21). Additionally, people can have diverse perceptions of their health influenced by cultural contexts, social positions, and personal health experiences (e.g. people suffering from the same illness for a longer time may report better levels of health due to various coping and self-management strategies) (22, 44). Nevertheless, self-reported health seems to be an important measure which can capture components of health or vulnerability that other measures alone cannot.

Strengths, limitations and future research

The availability of nationwide data on a wide range of risk and protective factors to vulnerability in many different domains was an important strength of this study. The outcome 'multidimensional vulnerability' was also based on 42 variables (12). Additionally, we conducted several sensitivity analyses, all of which yielded similar results, underscoring the robustness of our model. However, this study also had several limitations, mostly related to the data. One limitation concerns the representativeness of the dataset used to construct and evaluate the predictive models. It is possible that some factors (e.g. asylum seeker status) did not emerge as primary predictors because they were less present among the 4172 women, despite their association with vulnerability and adverse outcomes in the literature (45, 46). This may have also led to a slight underestimation of the actual percentage of multidimensional vulnerability. Additionally, we missed data on important topics that can contribute to vulnerability such as stress, health literacy, coping skills, and adverse (childhood) experiences including violence. Another limitation is that we insufficiently considered the dynamics around pregnancy in relation to vulnerability, since we merely incorporated data prior to pregnancy that can be subject to change. Future research should take into account that vulnerability can exist prior to pregnancy, but also

arise or change during pregnancy, childbirth or after birth (9). Also a consideration of the role of the father or woman's partner and wider social network could contribute to more insights into vulnerability and better predictions.

CONCLUSIONS

This study shows that it is feasible to predict multidimensional vulnerability at population-level using solely routinely collected data. Routinely collected data is readily available for the entire population, thereby providing a robust foundation for longitudinal monitoring and policy formulation at population-level. Nevertheless, while predictions are fairly accurate, adding self-reported data is of added value.

Funding
This study was supported by funding. The Dutch Ministry of Health, Welfare and Sport funded the monitor of the Dutch Solid Start program that was conducted by the National Institute for Public Health and the Environment (Joyce Molenaar, Ka Yin Leung, Peter Paul Klein, Jeroen Struijs, project numbers V/010038/01 and V/060438/22). The Bernard van Leer Foundation provided funding for different projects on the topic of vulnerability before, during and after pregnancy for Erasmus MC (Lindsey van der Meer, grant number NET-2017-096). Jessica Kiefte–de Jong received subsidy from the Netherlands Organization of Health Research and Development to study vulnerable and unintended pregnancies (grant number 554002006). The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing the manuscript and submission of the article for publication.

Conflicts of interests: None declared

Data availability
We are unable to share the individual data used for this study as data linkage and analysis was conducted within the highly safeguarded Remote Access (RA) platform of Statistics Netherlands (25). All data within this platform are pseudonymised to ensure data safety and confidentiality. Access to the data from Perined, Statistics Netherlands, and the Public Health Monitor 2016 can be requested from the relevant parties.

Acknowledgements
We thank Maarten Schipper and Albert Wong for valuable discussions regarding the statistical analyses. Additionally, we thank Agatha van Meijeren and Carolien van den Brink for providing feedback to a draft version of the manuscript.

Key points

- Considering the combination of both social risk and protective factors related to vulnerability is needed to identify pregnant women at risk of adverse outcomes
- Multidimensional vulnerability at population-level can be predicted using solely routinely collected data
- Using self-reported data in addition to routinely collected data can be relevant to further improve the prediction of multidimensional vulnerability
- The strongest predictors to multidimensional vulnerability are related to self-reported health, socioeconomic characteristics and healthcare expenditures and utilization
- Without additional data collection, routinely collected data could provide insight in the prevalence, geographical distribution and trends in multidimensional vulnerability at population-level, which can be used for longitudinal monitoring and the formulation of policies.

REFERENCES

1. Barker DJ. The origins of the developmental origins theory. J Intern Med. 2007;261(5):412-7.
2. National Academies of Sciences, Engineering and Medicine. Vibrant and healthy kids: Aligning science, practice, and policy to advance health equity. Washington, DC: The National Academies Press; 2019.
3. Arcaya MC, Arcaya AL, Subramanian SV. Inequalities in health: definitions, concepts, and theories. Global health action. 2015;8(1):27106.
4. Ministry of Health, Welfare and Sport. Solid Start - the action programme. The Hague; 2020. Available from: https://www.government.nl/documents/publications/2020/08/24/solid-start-the-action-programme
5. de Groot N, Bonsel GJ, Birnie E, Valentine NB. Towards a universal concept of vulnerability: broadening the evidence from the elderly to perinatal health using a Delphi approach. PLoS One. 2019;14(2):e0212633.
6. Colciago E, Merazzi B, Panzeri M, Fumagalli S, Nespoli A. Women's vulnerability within the childbearing continuum: A scoping review. European Journal of Midwifery. 2020;4.
7. van der Meer L, Ernst H, Blanchette L, Steegers E. Een kwetsbare zwangere, wat is dat eigenlijk? [A vulnerable pregnant woman, what is that actually?]. Medisch Contact. 2020.
8. Spini D, Bernardi L, Oris M. Toward a life course framework for studying vulnerability. Research in Human Development. 2017;14(1):5-25.
9. Scheele J, van der Vliet–Torij HH, Wingelaar-Loomans E, Goumans M. Defining vulnerability in European pregnant women, a Delphi study. Midwifery. 2020;86:102708.
10. Molenaar JM, Boesveld IC, Kiefte-de Jong JC, Struijs JN. Monitoring the Dutch Solid Start Program: Developing an Indicator Set for Municipalities to Monitor their First Thousand Days-Approach. International Journal of Integrated Care. 2022;22(4).
11. van der Meer L, Waelput AJ, Steegers EA, Bertens LC. Creating a sense of urgency and provoking action–An example on the use of heat maps to address perinatal health inequalities. Preventive Medicine Reports. 2022;30:102058.
12. Molenaar J, van der Meer L, Bertens L, de Vries E, Waelput A, Knight M, et al. Defining vulnerability subgroups among pregnant women using pre-pregnancy information: a latent class analysis. Eur J Public Health. 2023;33(1):25-34.
13. Struijs JN, Drewes HW, Heijink R, Baan CA. How to evaluate population management? Transforming the Care Continuum Alliance population health guide toward a broadly applicable analytical framework. Health Policy. 2015;119(4):522-9.
14. Care Continuum Alliance (CCA). Implementation and evaluation: a population health guide for primary care models. Washington, DC; 2021.
15. Girwar SAM, Jabroer R, Fiocco M, Sutch SP, Numans ME, Bruijnzeels MA. A systematic review of risk stratification tools internationally used in primary care settings. Health Science Reports. 2021;4(3):e329.
16. Ardesch F, Meulendijk M, Kist J, Vos R, Vos H, Kiefte-de Jong J, et al. The introduction of a data-driven population health management approach in the Netherlands since 2019: The Extramural LUMC Academic Network data infrastructure. Health Policy. 2023;132(104769).
17. Scheefhals ZT, de Vries EF, Molenaar JM, Numans ME, Struijs JN. Observational Data for Integrated Maternity Care: Experiences with a Data-Infrastructure for Parents and Children in the Netherlands. International Journal of Integrated Care. 2023;23(4).
18. Fayers PM, Sprangers MA. Understanding self-rated health. The Lancet. 2002;359(9302):187-8.
19. Kuipers YJF, Mestdagh E. The experiential knowledge of migrant women about vulnerability during pregnancy: A woman-centred mixed-methods study. Women and Birth. 2022;35(1):70-9.

20. Miilunpalo S, Vuori I, Oja P, Pasanen M, Urponen H. Self-rated health status as a health measure: the predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. J Clin Epidemiol. 1997;50(5):517-28.
21. Au N, Johnston DW. Self-assessed health: what does it mean and what does it hide? Soc Sci Med. 2014;121:21-8.
22. Paul P, Nguemdjo U, Kovtun N, Ventelou B. Does self-assessed health reflect the true health state? Int J Environ Res Public Health. 2021;18(21):11153.
23. Perined. Over Perined [about Perined] n.d. [10/17/2023]. Available from: https://www.perined.nl/over-perined
24. Bakker BF, Van Rooijen J, Van Toor L. The system of social statistical datasets of Statistics Netherlands: An integral approach to the production of register-based social statistics. Statistical Journal of the IAOS. 2014;30(4):411-24.
25. Statistics Netherlands (CBS). About CBS n.d. [Date cited: 10/17/2023]. Available from: https://www.cbs.nl/en-gb/about-us/
26. Community Health Services Statistics Netherlands and the National Institute for Public Health and the Environment. Public Health Monitor 2012, 2016 and 2020. Available from: https://www.monitorgezondheid.nl/gezondheidsmonitor-volwassenen-en-ouderen
27. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of statistical software. 2011;45:1-67.
28. Breiman L. Random forests. Machine learning. 2001;45:5-32.
29. Davis J, Goadrich M, editors. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning; 2006.
30. Wright MN, Wager S, Probst P. Package 'ranger' 2023. Available from: https://mirror.las.iastate.edu/CRAN/web/packages/ranger/ranger.pdf
31. Parvandeh S, Yeh H-W, Paulus MP, McKinney BA. Consensus features nested cross-validation. Bioinformatics. 2020;36(10):3093-8.
32. Janitza S, Celik E, Boulesteix A-L. A computationally fast variable importance test for random forests for high-dimensional data. Advances in Data Analysis and Classification. 2018;12:885-915.
33. van Blarikom E, de Kok B, Bijma HH. "Who am I to say?" Dutch care providers' evaluation of psychosocial vulnerability in pregnant women. Soc Sci Med. 2022;307:115181.
34. Luna F. Identifying and evaluating layers of vulnerability–a way forward. Dev World Bioeth. 2019;19(2):86-95.
35. Tully KP, Gibson AN, Pearsall MS, Umstead K, Gill C, Stuebe AM. Screening and referral for social determinants of health: Maternity patient and health care team perspectives. Health Equity. 2022;6(1):887-97.
36. Vezzoli M, Valtorta RR, Mari S, Durante F, Volpato C. Effects of objective and subjective indicators of economic inequality on subjective well-being: Underlying mechanisms. J Appl Soc Psychol. 2023;53(2):85-100.
37. Tan JJ, Kraus MW, Carpenter NC, Adler NE. The association between objective and subjective socioeconomic status and subjective well-being: A meta-analytic review. Psychol Bull. 2020;146(11):970.
38. Navarro-Carrillo G, Alonso-Ferres M, Moya M, Valor-Segura I. Socioeconomic status and psychological well-being: Revisiting the role of subjective socioeconomic status. Front Psychol. 2020;11:1303.
39. Cundiff JM, Matthews KA. Is subjective social status a unique correlate of physical health? A meta-analysis. Health Psychol. 2017;36(12):1109.
40. Haushofer J, Fehr E. On the psychology of poverty. Science. 2014;344(6186):862-7.
41. Cundiff JM, Boylan JM, Muscatell KA. The pathway from social status to physical health: Taking a closer look at stress as a mediator. Curr Dir Psychol Sci. 2020;29(2):147-53.

42. Adamkovič M, Martončik M. A review of consequences of poverty on economic decision-making: A hypothesized model of a cognitive mechanism. Front Psychol. 2017;8:1784.
43. De Bruijn E-J, Antonides G. Poverty and economic decision making: a review of scarcity theory. Theory and Decision. 2022;92(1):5-37.
44. Roudijk B, Donders R, Stalmeier P. Cultural values: can they explain self-reported health? Qual Life Res. 2017;26:1531-9.
45. Heslehurst N, Brown H, Pemu A, Coleman H, Rankin J. Perinatal health outcomes and care among asylum seekers and refugees: a systematic review of systematic reviews. BMC Med. 2018;16(1):1-25.
46. Tankink JB, Verschuuren AE, Postma IR, van der Lans PJ, de Graaf JP, Stekelenburg J, Mesman AW. Childbirths and the Prevalence of Potential Risk Factors for Adverse Perinatal Outcomes among Asylum Seekers in The Netherlands: A Five-Year Cross-Sectional Study. Int J Environ Res Public Health. 2021;18(24):12933.

APPENDIX 1.
Methodology

Sensitivity analyses: XGBoost and Lasso regression

XGBoost (extreme gradient boosting) is a machine learning technique that iteratively builds multiple shallow decision trees (1). Similar to RF, it is a flexible algorithm without assuming a functional form. Logistic regression, on the other hand, does assume a strong functional form, i.e. a linear relation between the independent variables and log odds. Logistic regression is a standard approach for binary classification with a long history in literature. The logistic regression analysis was conducted with lasso penalty to shrink coefficients towards zero such that less important variables are left out the model (2).

Nested cross-validation

The three techniques RF, Lasso and XGBoost each have their own set of hyperparameters that need to be chosen for the models. For RF, the default hyperparameter settings in the R-package 'ranger' (3) were used, as these default settings generally yield good performance. The parameter to choose for Lasso (R-package 'glmnet' (4)) was the lambda, which defines the penalty, and for XGboost (R-package 'xgboost' (5)) the number of trees and tree-depth. For Lasso and XGBoost we used cross-validation to choose the hyperparameters. In addition, as the models predict the probability of multidimensional vulnerability, we need to choose the threshold at which all predicted probabilities above that threshold are classified as multidimensional vulnerable 'yes' (and as 'no' below that threshold). To choose the hyperparameters and threshold probability, and finally to assess the performance of the models, we used nested-cross validation. Firstly, the dataset of 4172 women was split into six folds: 5 parts train-set, 1 part test-set (outer loop). Secondly, in the nesting step (inner loop), each train set from the outer loop was again split into five folds: 4 parts train-set, 1 part validation-set. During each split, we made sure that the percentage of multidimensional vulnerability was approximately equal in each part.

Firstly, using the cross-validation of the inner-loop, we chose the hyperparameters: for Lasso, we chose the average lambda across the five validation folds and for XGBoost we selected the hyperparameters for which the average AUC over the five folds was highest. Secondly, using the defined hyperparameters, and the same inner loop, we selected the threshold probability that yielded the highest F1-value on the validation set (averaged over five validation folds). Thirdly, by utilizing both the selected hyperparameters and average optimal thresholds, we calculated the F1-value on the test set of the outer loop that has not been used in selecting hyperparameters and threshold. This three-step process was repeated for the 6 folds of the outer loop, yielding the F1-measure of the model averaged over the 6 test-sets. Using the average optimal thresholds (as well as the defined hyperparameters), we fitted the model one last time on all data. The final model can be utilized for predicting outcomes on new datasets. We used the final RF-model in our next analyses.

Variable importance with OOB-observations, including sensitivity analyses

Variable importance was measured with RF in the following way. RF takes a bootstrap sample for every tree that it constructs. The data that are not used in the bootstrap sample are called the out-of-bag (OOB) observations. RF makes a prediction for these OOB-observation based on the tree that is constructed on the bootstrap-sample, leading to an OOB-error. Next, to determine the tree-specific importance of a variable, a variable is randomly shuffled (permuted) in the bootstrap sample. In this new variant of the bootstrap sample, a new tree is grown which gives a new OOB-error. This OOB-error is then compared to the original OOB-error. If permuting a variable increases the error, it is considered important as the model relied on it for prediction. Consequently, by permuting a variable and comparing the OOB-error rates of the predictions before and after permutation (6), we obtain a measure of variable importance for each variable for a single tree. The OOB-errors increase for each variable are averaged over all trees and compared. The average of all these tree importance values yields the ranking of variables for the model (7).

As sensitivity analyses, we also checked the permutation importance and Partial Dependence Plots (PDP) (8-10). Permutation importance permutes the values of a specific variable in the full dataset (rather than individual trees) to measure the impact on the model's performance. The PDP of each variable provides insight into the direction and strength of the relationship with the dependent variable while holding all other predictors constant. We checked whether the direction of the important variables aligned with their categorization as risk or protective factors.

References

1. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016;. p. 785-94.
2. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. Journal of statistical software 2010; 33(1): 1.
3. Wright MN, Wager S, Probst P. Package 'ranger'. 2023. Available from: https://mirror.las.iastate.edu/CRAN/web/packages/ranger/ranger.pdf.
4. Friedman J, Hastie T, Tibshirani R, Narasimhan B, Tay K, Simon N, Qian J. Package 'glmnet'. 2023.
5. Chen T, He T, Benesty M, Khotilovich V. Package 'xgboost'. 2023. Available from: https://cran.utstat.utoronto.ca/web/packages/xgboost/xgboost.pdf.
6. Breiman L. Random forests. Machine learning 2001; 45: 5-32.
7. Janitza S, Celik E, Boulesteix A-L. A computationally fast variable importance test for random forests for high-dimensional data. Advances in Data Analysis and Classification 2018; 12: 885-915.
8. Molnar C. Interpretable machine learning: Lulu. com; 2020.
9. Fisher A, Rudin C, Dominici F. All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. J Mach Learn Res 2019; 20(177): 1-81.
10. Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of statistics 2001: 1189-232.

APPENDIX 2.
Results

Table 1. Characteristics of the study populations (including missing data)

		Original study population *(objective 1, 2 and 3)*	**All unique pregnancies 2017 – 2021** *(objective 1)*
	Total n =	4172	807.904
		n (%)	n (%)
Individual characteristics			
Age	19-23	306 (7.3)	43797 (5.4)
	24-35	3528 (84.6)	624292 (77.3)
	>35	338 (8.1)	135882 (16.8)
	Missing	-	*3933 (0.5)*
Ethnicity	Non-Western	420 (10.1)	170968 (21.2)
	Western	343 (8.2)	89970 (11.1)
	Native Dutch	3409 (81.7)	546624 (67.7)
	Missing	-	*342 (<0.1)*
Parity[a]	Nullipara	1755 (42.1)	355979 (44.1)
	Primipara, multipara	2410 (57.8)	449416 (55.6)
	Missing	*<10 (<0.2)*	*2509 (0.3)*
Asylum seeker status	Yes	39 (0.9)	19582 (2.4)
	No	4133 (99.1)	788322 (97.6)
Socioeconomic characteristics			
Educational level	Low	328 (7.9)	80849 (10.0)
	Moderate	1513 (36.3)	289681 (35.9)
	High	2303 (55.2)	340669 (42.2)
	Missing	*28 (0.7)*	*96705 (12.0)*
Household income	Low	202 (4.8)	60682 (7.5)
	Moderate	3348 (80.2)	613979 (76.0)
	High	591 (14.2)	103122 (12.8)
	Missing	*31 (0.7)*	*30121 (3.7)*
Socioeconomic position (occupational status)	No income/receiving benefits	532 (12.8)	132780 (16.4)
	Student	82 (2.0)	24988 (3.1)
	Paid work	3502 (83.9)	632685 (78.3)
	Missing	*56 (1.3)*	*17451 (2.2)*
Debts and payment arrears	Yes	45 (1.1)	19357 (2.4)
	No	4127 (98.9)	788547 (97.6)
Insufficient financial resources	Yes	524 (12.6)	
	No	3267 (78.3)	
	Missing	*381 (9.1)*	
Permanent contract	No	1929 (46.2)	433393 (53.6)
	Yes	2243 (53.8)	374511 (46.4)
Full-time contract	No	1925 (46.1)	653069 (80.8)
	Yes	2247 (53.9)	154835 (19.2)
Lifestyle factors			
Smoking	Yes	661 (15.8)	
	No	3315 (79.5)	
	Missing	*196 (4.7)*	
Alcohol use	Yes (excessive)	418 (10.0)	
	No	3503 (84.0)	
	Missing	*251 (6.0)*	
Physical activity	Less than recommended	1696 (40.7)	
	As recommended or more	2158 (51.7)	
	Missing	*318 (7.6)*	

Table 1. Continued.

		Original study population *(objective 1, 2 and 3)*	**All unique pregnancies 2017 – 2021** *(objective 1)*
	Total n =	4172	807.904
		n (%)	n (%)
BMI	Unhealthy BMI Healthy BMI *Missing*	1386 (33.2) 2641 (63.3) *145 (3.5)*	
Household characteristics			
Type of household	One-person/parent household Other	353 (8.5) 3819 (91.5)	123697 (15.3) 684207 (84.7)
Marital status	Unmarried Married	2147 (51.5) 2025 (48.5)	433383 (53.6) 374521 (46.4)
Dissolution of marriage	Yes No	58 (1.4) 4114 (98.6)	14928 (1.8) 792976 (98.2)
Household size	≥6 persons <6 persons *Missing*	93 (2.2) 4079 (97.8) -	20412 (2.5) 757300 (93.7) *30192 (3.7)*
Youth support utilization	Yes No	102 (2.4) 4070 (97.6)	79473 (9.8) 728431 (90.2)
Self-reported health			
Perceived health status	Negative Positive *Missing*	465 (11.1) 3653 (87.6) *54 (1.3)*	
Long-term illness	Yes No *Missing*	747 (17.9) 3362 (80.6) *63 (1.5)*	
Restricted by health	Yes No *Missing*	724 (17.4) 3330 (79.8) *118 (2.8)*	
Healthcare expenditures and utilization			
Overall healthcare expenditures	High Low-average *Missing*	824 (19.8) 3297 (79.0) *51 (1.2)*	151443 (18.7) 627116 (77.6) *29345 (3.6)*
GP expenditures	High Low-average *Missing*	827 (19.8) 3308 (79.3) *37 (0.9)*	145166 (18.0) 633247 (78.4) *29491 (3.7)*
Hospital expenditures	High Low or none *Missing*	413 (9.9) 3708 (88.9) *51 (1.2)*	76523 (9.5) 731381 (90.5) *0 (0.0)*
Medication use	High Low or none	428 (10.3) 3744 (89.7)	92295 (11.4) 715609 (88.6)
Addiction related care utilization	Yes No	23 (0.6) 4149 (99.4)	4466 (0.6) 803438 (99.4)
Psychosocial characteristics			
Mental healthcare utilization	Yes No *Missing*	228 (5.5) 3907 (93.6) *37 (0.9)*	50630 (6.3) 739093 (91.5) *18181 (2.3)*
Risk of depression or anxiety disorders	Moderate – high risk No or low risk *Missing*	1716 (41.1) 2256 (54.1) *200 (4.8)*	
Loneliness	Feeling lonely Not feeling lonely *Missing*	1100 (26.4) 2719 (65.2) *353 (8.5)*	

Table 1. Continued.

		Original study population *(objective 1, 2 and 3)*	**All unique pregnancies 2017 – 2021** *(objective 1)*
	Total n =	4172	807.904
		n (%)	n (%)
Feelings of control over life	Low	144 (3.5)	
	Moderate	2741 (65.7)	
	High	1006 (24.1)	
	Missing	*281 (6.7)*	
Mild intellectual disability	Yes	13 (0.3)	7187 (0.9)
	No	4159 (99.7)	800717 (99.1)
Life-events			
Crime suspect	Yes	95 (2.3)	35393 (4.4)
	No	4077 (97.7)	772511 (95.6)
Crime victim	Yes	874 (20.9)	171564 (21.2)
	No	3298 (79.1)	636340 (78.8)
Having been detained[a]	Yes	not shown	2592 (0.3)
	No	not shown	805312 (99.7)
History of frequent moving[b]	Yes	53 (1.3)	14081 (1.7)
	No	4119 (98.7)	793823 (98.3)
Loss of a family member	Yes	147 (3.5)	32839 (4.1)
	No	4025 (96.5)	775065 (95.9)
Living conditions			
Home ownership	Rented	990 (23.7)	264519 (32.7)
	Owner occupied	3099 (74.3)	531526 (65.8)
	Missing	*83 (2.0)*	*11859 (1.5)*
Motorized vehicle ownership	No	494 (11.8)	162579 (20.1)
	Yes	3678 (88.2)	645325 (79.9)
Proximity to General Practitioners' (GP) office	>3 km	265 (6.4)	62040 (7.7)
	<3 km	3847 (92.2)	740876 (91.7)
	Missing	*60 (1.4)*	*4988 (0.6)*
Liveability neighbourhood	Low-mediocre	273 (6.5)	87013 (10.8)
	High	3695 (88.6)	714817 (88.5)
	Missing	*204 (4.9)*	*6074 (0.8)*

a: Following guidelines of Statistics Netherlands, the data of some variables were rounded (parity) or not shown (having been detained) to prevent the disclosure of information about individuals.
b: Erratum: in the original paper of Molenaar et al. (12), there were inaccuracies in the reported percentages of the variable 'history of frequent moving', which consequently deviate from the values presented here.
Missing data are shown in italic.

Objective 1

Table 2. Hyperparameters and thresholds used for RF, XGBoost and Lasso regression

	Hyperparameters and thresholds			
	Threshold (average optimal)	Lambda	Tree depth	Number of trees
Random Forest	0.37 (0.01)	NA	default	default
XGBoost	0.36 (0.04)	NA	2	51
Lasso regression	0.26 (0.04)	0.01 (0.003)	NA	NA

Results based on analyses among study population of 4172 women

Table 3. Results of the RF and sensitivity analyses for the six separate folds

		Metrics				
		AUC	F1-measure	Precision	Recall/ sensitivity	Specificity
Random Forest	**Average**	**0.98 (0.00)**	**0.70 (0.03)**	**0.74 (0.06)**	**0.66 (0.04)**	**0.98 (0.00)**
	Fold 1	0.98	0.68	0.69	0.67	0.98
	Fold 2	0.98	0.68	0.68	0.69	0.98
	Fold 3	0.98	0.70	0.74	0.67	0.98
	Fold 4	0.98	0.75	0.83	0.68	0.99
	Fold 5	0.98	0.72	0.74	0.69	0.98
	Fold 6	0.97	0.66	0.75	0.59	0.99
XGBoost	**Average**	**0.98 (0.00)**	**0.68 (0.04)**	**0.70 (0.02)**	**0.67 (0.08)**	**0.98 (0.00)**
	Fold 1	0.98	0.68	0.68	0.69	0.98
	Fold 2	0.98	0.68	0.68	0.69	0.98
	Fold 3	0.98	0.61	0.70	0.55	0.98
	Fold 4	0.98	0.75	0.72	0.77	0.98
	Fold 5	0.98	0.66	0.74	0.60	0.99
	Fold 6	0.97	0.70	0.71	0.70	0.98
Lasso regression	**Average**	**0.98 (0.01)**	**0.68 (0.04)**	**0.67 (0.07)**	**0.70 (0.07)**	**0.98 (0.01)**
	Fold 1	0.97	0.64	0.54	0.78	0.96
	Fold 2	0.98	0.71	0.65	0.77	0.98
	Fold 3	0.97	0.65	0.71	0.60	0.98
	Fold 4	0.98	0.74	0.74	0.73	0.98
	Fold 5	0.98	0.66	0.68	0.64	0.98
	Fold 6	0.97	0.69	0.70	0.67	0.98

Results based on analyses among study population of 4172 women

Table 4. Average proportion of multidimensional vulnerability in the Netherlands over the years 2017 – 2021 for the three regression models

		2017	2018	2019	2020	2021
RF	Average proportion	0.081	0.079	0.079	0.077	0.072
	SD	8.21E-05	2.65E-05	6.7E-05	5.93E-05	0.000105
XGBoost	Average proportion	0.090	0.088	0.088	0.085	0.080
	SD	0.000105	0.000151	9.21E-05	7.95E-05	5.55E-05
Lasso regression	Average proportion	0.100	0.098	0.098	0.097	0.091
	SD	0.000116	8.69E-05	7.05E-05	5.66E-05	9.52E-05

Results based on analyses among all unique pregnancies from 2017 – 2021 (n = 807.904)

Table 5. Complete cases

		2017	2018	2019	2020	2021
RF	Proportion	0.086	0.084	0.083	0.080	0.075
XGBoost	Proportion	0.095	0.092	0.091	0.087	0.081
Lasso regression	Proportion	0.106	0.103	0.102	0.099	0.093

Results based on analyses among all unique pregnancies from 2017 – 2021 with complete data on all variables

Objective 2

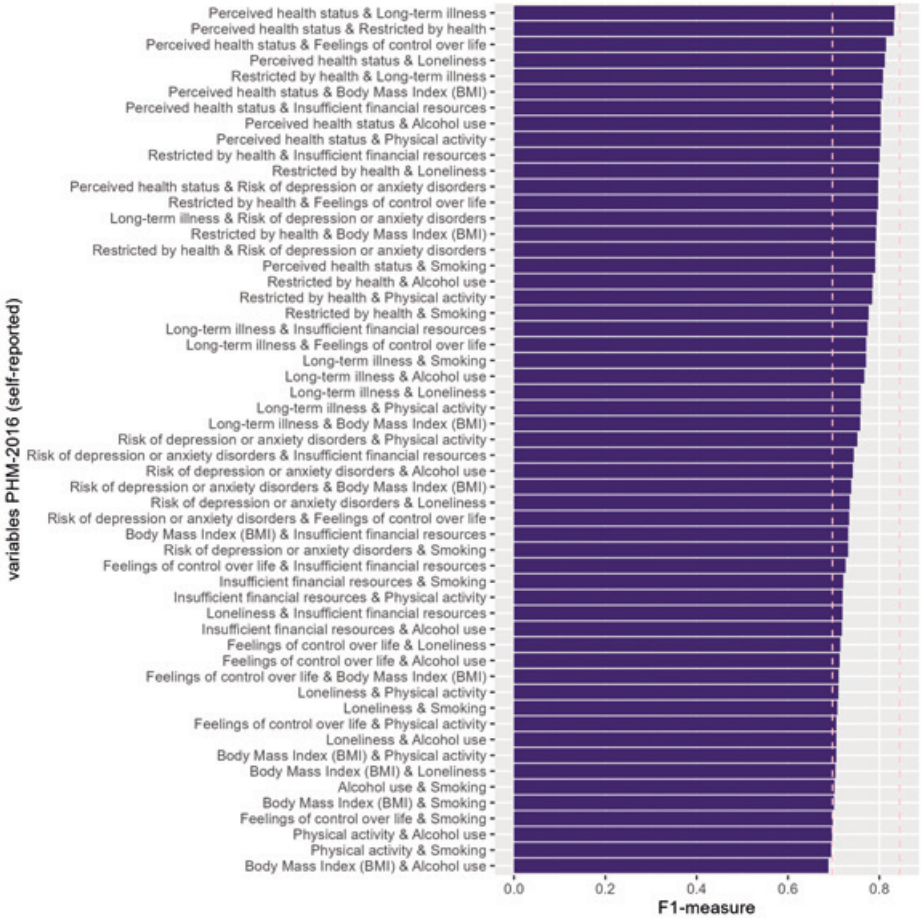


Figure 1. Two variables on self-reported health, wellbeing and lifestyle added to the RF-models with solely routinely collected data.
Results based on analyses among study population of 4172 women

Objective 3

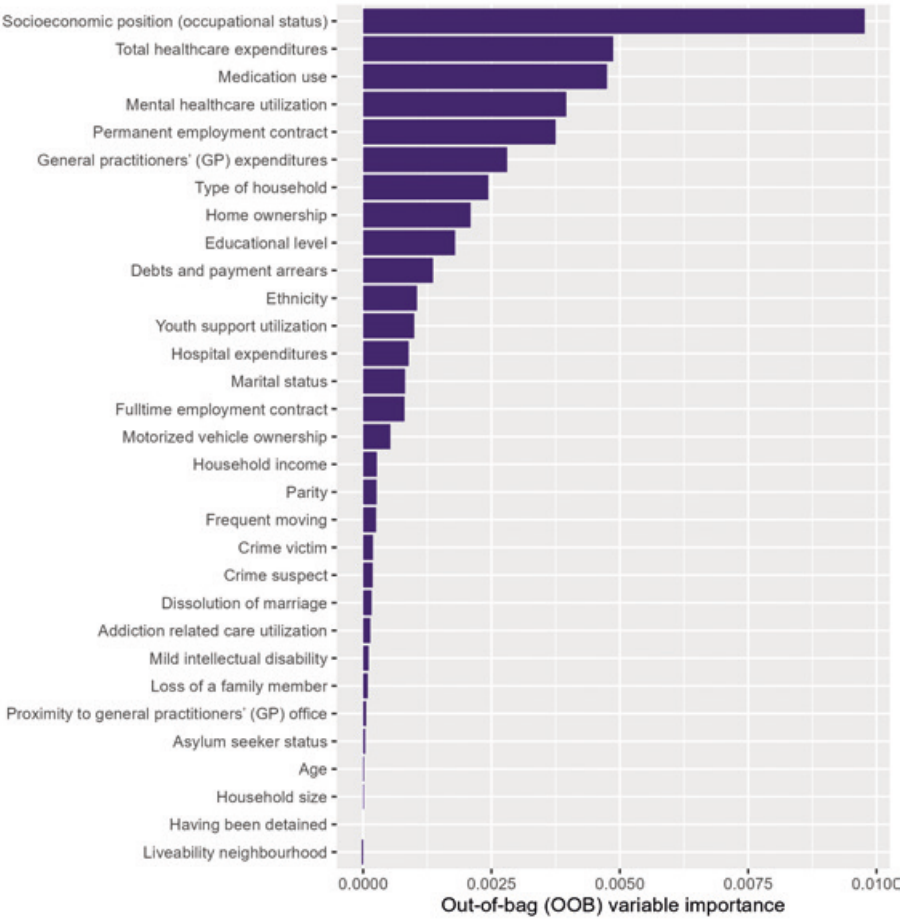


Figure 2. Variable importance ranking of the RF-model for 'multidimensional vulnerability', using 31 variables (solely routinely collected data).
Results based on analyses among study population of 4172 women

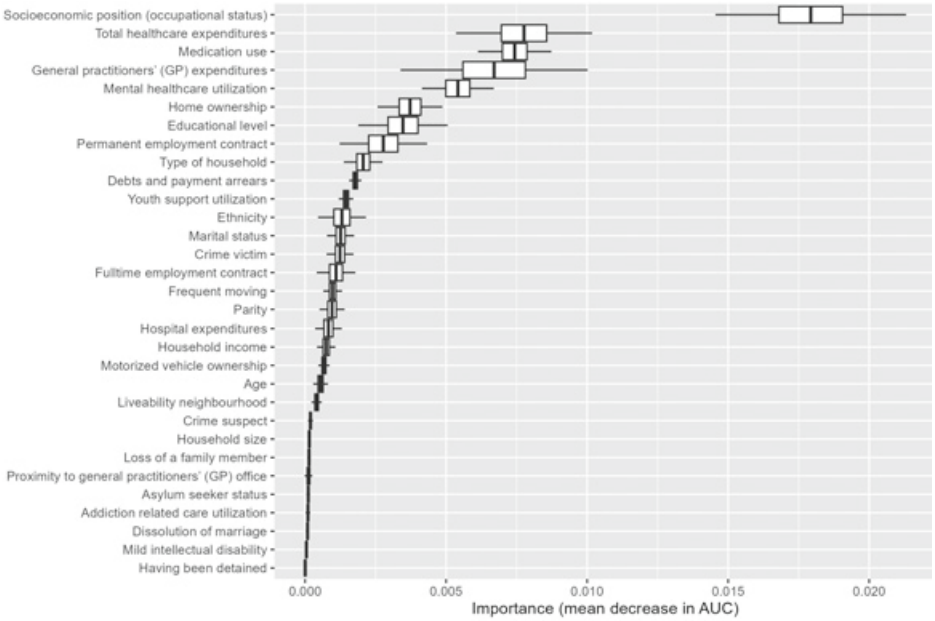


Figure 3. Permutation importance ranking of the RF-model for 'multidimensional vulnerability', using 31 variables (solely routinely collected data)
Results based on analyses among study population of 4172 women

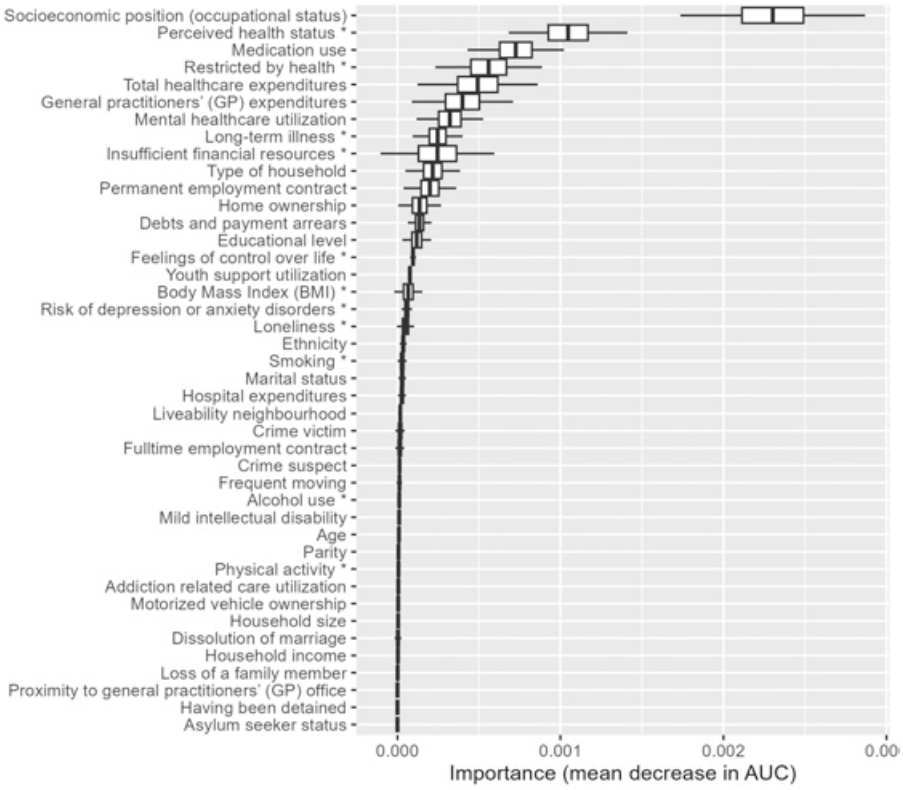


Figure 4. Permutation importance ranking of the RF-model for 'multidimensional vulnerability', using the set of 42 variables (both routinely collected data and self-reported data)
Results based on analyses among study population of 4172 women