

BMJ Open Development of a casemix classification to predict costs of home care in the Netherlands: a study protocol

Arianne Mathilda Josephus Elissen ,¹ Gertjan Sebastiaan Verhoeven,^{2,3} Maud Hortense de Korte,^{2,3} Anne Odilia Emile van den Bulck,¹ Silke Friederike Metzethin,⁴ Lieuwe Christiaan van der Weij,² Jaap Stam,² Dirk Ruwaard,¹ Misja Chiljon Mikkers^{2,3}

To cite: Elissen AMJ, Verhoeven GS, de Korte MH, *et al.* Development of a casemix classification to predict costs of home care in the Netherlands: a study protocol. *BMJ Open* 2020;**10**:e035683. doi:10.1136/bmjopen-2019-035683

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-035683>).

Received 11 November 2019
Revised 16 January 2020
Accepted 23 January 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Health Services Research, CAPHRI Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands

²Dutch Healthcare Authority, Utrecht, The Netherlands

³Department of Economics, Tilburg University, Tilburg, The Netherlands

⁴Health Services Research, Maastricht University, Maastricht, The Netherlands

Correspondence to

Dr Arianne Mathilda Josephus Elissen;
a.elissen@maastrichtuniversity.nl

ABSTRACT

Introduction Compared with fee-for-service systems, prospective payment based on casemix classification is thought to promote more efficient, needs-based care provision. We aim to develop a casemix classification to predict the costs of home care in the Netherlands. **Methods and analysis** The research is designed as a multicentre, cross-sectional cohort study using quantitative methods to identify the relative cost predictors of home care and combine these into a casemix classification, based on individual episodes of care. The dependent variable in the analyses is the cost of home care utilisation, which is operationalised through various measures of formal and informal care, weighted by the relative wage rates of staff categories. As independent variables, we will use data from a recently developed Casemix Short-Form questionnaire, combined with client information from participating home care providers' (nursing) classification systems and data on demographics and care category (ie, a classification mandated by health insurers). Cost predictors are identified using random forest variable importance measures, and then used to build regression tree models. The casemix classification will consist of the leaves of the (pruned) regression tree. Internal validation is addressed by using cross-validation at various stages of the modelling pathways. The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement was used to prepare this study protocol.

Ethics and dissemination The study was classified by an accredited Medical Research Ethics Committee as not subject to the Dutch Medical Research Involving Human Subjects Act. Findings are expected in 2020 and will serve as input for the development of a new payment system for home care in the Netherlands, to be implemented at the discretion of the Dutch Ministry of Health, Welfare and Sports. The results will also be published in peer-reviewed publications and policy briefs, and presented at (inter) national conferences.

INTRODUCTION

Casemix classification is defined as 'the act of grouping healthcare cases into clinically similar groups that are believed to also consume a similar basket of resources and,

Strengths and limitations of this study

- In developing a casemix classification for home care, the choice of potential casemix predictors is based on previous studies reported in the international scientific literature.
- Data necessary to operationalise the dependent and independent variables of interest are collected using a recently developed Casemix Short-Form questionnaire, and from participating home care providers' electronic health records and administrative databases.
- The 'Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis' statement is used as a guide when reporting on this study.
- Further research is needed for the purpose of external validation of the casemix classification using a different data set from different providers.
- Since this study is conducted in the Netherlands, further investigation and refinement may be needed before the casemix classification can be applied in other countries.

by extension, have similar costs'.¹ It is an important tool for health service reimbursement, particularly within prospective, per-case payment systems. Under such systems, care providers bear a certain financial risk because their costs for a given case can be higher than the ex ante defined reimbursement.² This should incentivise provider efficiency, in particularly compared with fee-for-service (FFS) systems, where care providers are paid for each item of service they provide. However, there is a risk that providers may attempt to reduce costs by, for example, providing too little care or only accepting cases that are profitable under the reimbursement scheme.^{2 3} Such negative effects are more likely to occur when there is a weak relationship between the prospective payment and expected cost of care, and can



ultimately be detrimental to patients and lead to higher macrolevel care costs.² By aligning providers' level of reimbursement with their expected costs, casemix classification can reduce incentives for undesirable strategic behaviour, although monitoring quality of care is equally important.^{4,5}

Since the 1970s, casemix classifications have been developed for many healthcare sectors, including hospital care (eg, the diagnosis-related groups),⁶ nursing home care (eg, resource utilisation groups-III),⁷ inpatient psychiatric care (eg, the psychiatric diagnostic groupings)⁸ and ambulatory care (eg, ambulatory care groups).⁹ Home care is arguably one of the more challenging sectors for casemix classification, particularly compared with inpatient care. As early as 1987, Manton and Hausner noted that 'a casemix measure for community-based long-term care services is intrinsically more complex than that for acute care because it must describe a multidimensional system of health, functional and social needs evolving over a potentially long time span'.¹⁰ Indeed, the determinants of the need for home care include not only clients' medical diagnoses, but also their physical and cognitive functioning.¹¹⁻¹³ In addition, there are challenges in defining home care episodes, which can range from days to years. Clients' living arrangements, family structure and social network are further concerns, since home care funding should not disincentivise the provision of informal care.^{10,13,14} Despite these complexities, a number of casemix classifications for prospective payment of home care exist: notable examples include the Home and Community Services Support (HCSS) model used in New Zealand⁴ and the Home Health Resource Groupings model from the USA.^{15,16}

This study is part of a project in which the Dutch Healthcare Authority, in collaboration with academic partners, is developing a prospective, per-case payment system for home care in the Netherlands. A casemix classification should form the basis for the new system. Although adopting an existing classification from elsewhere has advantages, particularly in terms of the time and money required, the development of a unique classification allows it to be tailored to local policy goals and context.¹⁷ Important policy goals in reforming the Dutch payment system for home care are to incentivise value rather than volume, and to serve the needs of all home care beneficiaries better than the current fee-per-hour (ie, FFS) system. Using standardised (nursing) classification data for casemix classification offers plenty of potential for achieving these policy goals, without adding a considerable administrative burden to home care providers. Previous research suggests that such data, which are routinely registered in home care and stored in providers' electronic health records (EHRs), can be used to predict casemix efficiently.¹⁸ However, there are also barriers that impede immediate adoption, including the use of multiple classification systems in Dutch home care (eg, NANDA-I, Omaha or InterRAI) and variations in registration practices.¹⁹ Moreover, recent studies suggest

that additional client information, beyond that available in EHRs, could further improve the accuracy of casemix classification in home care.^{4,20}

For these reasons, the objectives of this study are to: (1) identify the relative cost predictors of home care services using data from a recently developed Casemix Short-Form (CM-SF) questionnaire, combined with EHR data and (2) based on these insights, develop a casemix classification for prospective, per-case-based payment of home care services. Regarding the latter, the aim is specifically to develop a scientifically robust classification with maximum predictive power: analysing the feasibility of the classification for policy and practice is not part of the scope of this study.

METHODS

We followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement to prepare this study protocol.²¹

Study design and source of data

The research is designed as a multicentre, cross-sectional cohort study using quantitative methods to identify the relative cost predictors for home care and combine these into a casemix classification, based on individual episodes of care. An 'episode of care' starts when a client: (1) receives an initial needs assessment for home care or (2) is formally reassessed as part of ongoing care, which is every 6 months in standard practice but can occur more frequently in the case of a care plan re-evaluation (due to changes in the client's status). An episode of care ends after a fixed period of 4 or 13 weeks. This a priori design choice was made to enable analysis of the effect of care episode duration on the accuracy of casemix prediction. Based on the findings, an informed decision can be made on the most appropriate time frame for prospective funding of Dutch home care, that is, per month (4 weeks) or per quarter (13 weeks). Clients who, during a 4-week or 13-week care episode, are no longer in need of home care from the current provider (eg, due to improved health status, relocation or death) are retained in the data sample.

The study is being conducted between June 2019 and March 2020: clients from four Dutch home care providers will be included from the start of the study until December 2019. Providers were selected based on their involvement in a funding experiment organised by health insurers focusing on alternatives to FFS, which started in 2016. They operate in various regions of the Netherlands and provide services to relatively large client populations (ie, between 2000 and 4000 clients per provider at any given time).

Participants

Study setting

Following recent work by Maurits,²² we define home care in the Dutch context as 'formal nursing services and

personal care provided by nursing staff in clients' own homes'. More specifically:

'Nursing services can be of a technical, supportive, rehabilitative or preventive nature. Personal care services relate to assistance with activities that are part of daily living, such as dressing, feeding and washing. Different types of care can be delivered to various types of patients, such as the chronically ill, disabled people, elderly people and people at the end of life. Home care encompasses both long-term care at home and short-term care at home, for instance, after discharge from the hospital'.²²

Home care in the Netherlands is provided mainly by: (1) registered nurses with either an associate degree (ie, senior secondary vocational education; European Qualifications Framework (EQF) educational level 4) or a Bachelor's degree (ie, University of Applied Sciences; EQF educational level 6); (2) certified nursing assistants (EQF educational level 3) and (3) care assistants. Although clear demarcations between the tasks of different nursing staff categories are lacking, tasks that are more complex (eg, wound care) are generally performed by registered nurses.²² In 2017, approximately 2040 home care providers (including self-employed nurses) provided services to more than 550 000 clients in the Netherlands.²³

Since 2015, the Dutch Health Insurance Act (HIA) obliges residents of the Netherlands to purchase a basic health insurance package for essential curative services. The HIA covers home care for clients who need care for less than 24 hours per day.²⁴ Registered nurses with a bachelor's degree (EQF level 6) are responsible for performing the formal assessment of care needs for services covered by the HIA, taking into account the self-reliance of citizens and the resources available in their social network. Needs assessment is supported by a (nursing) classification system, such as NANDA-I or Omaha, and is also used by nurses to draw up a care plan that includes aims, interventions and outcomes.^{22 24} Until recently, nurses had to evaluate the care plan with the client and/or informal caregiver every 6 months; now, however, they are free to determine when a re-evaluation takes place depending on the client and the care aims.

Eligibility criteria

Clients receiving home care services from any of the participating home care providers are eligible for this study, if they: (1) receive a formal needs assessment—either a first assessment or a reassessment—as part of regular home care procedures during the inclusion period, regardless of primary diagnosis and (2) subsequently receive home care services covered under the Dutch HIA. No further inclusion or exclusion criteria are applied.

Dependent variable (outcome)

The dependent variable in the analyses is the cost of home care utilisation. Following the approach of Björkgren *et al*,²⁵ we operationalise home care costs as various measures of formal and informal care time, weighted by the relative wage rates of the staff categories involved.

Table 1 Cost weights for formal care time per discipline and for informal care

Home care discipline	Cost per hour (in euros)	Cost weight
Care assistant	43	0.86
Certified nursing assistant (EQF 3)	48–52	0.96–1.04
Registered nurse (EQF 4)	61	1.22
Registered nurse (EQF 6)	77	1.54
Informal care	–	0.5

For the purpose of cost weighting, the average hourly wage rate for certified nursing assistants (€50) was set to 1.0. EQF, European Qualifications Framework.

Formal care cost

To estimate formal care cost, we collect data on the estimated total time (in hours) spent providing home care to a given client by the care providers' staff members. Formal care time is limited to services covered by the Dutch HIA, that is, nursing care and personal care services as defined by Maurits.²² Typically, time spent on a client is registered in various ways, including automatic generation using an electronic touch device in the client's home or manual registration of the preallocated or actual time in the EHR. When a provider uses multiple forms of time registration, we will choose the data source that most closely approximates the actual time spent on the client, in consultation with the provider.

Formal care hours are differentiated by the main staff categories involved in service provision, that is, registered nurses (EQF levels 4 or 6), certified nursing assistants (EQF level 3) and care assistants. To operationalise cost measures, formal care time per staff category is weighted by cost using hourly wage rates. Table 1 provides relative wage weights per staff category, based on a recent Dutch costing study.²⁶ The weights are standardised by setting the average rate for certified nursing assistants to 1.0.

Two dependent variables are constructed based on these data: mean formal care cost per client per week during the 4-week and 13-week periods following the needs assessment. Mean weekly costs are calculated by dividing the total of cost-weighted care hours by the episode of care (ie, 4 or 13 weeks). The casemix classification is intended to mitigate providers' financial risk under prospective, per-case payment due to different resource utilisation needs in home care. Predicting the mean total cost of formal care per client for each of the care episodes studied fits this purpose best. This means that no adjustments are made for clients whose service use ends prior to the conclusion of the fixed care episodes included in the analyses.

Informal care cost

Informal care time reflects the total amount of unpaid care provided by informal caregivers, such as family

members, friends and neighbours, and is defined as follows²⁷:

‘Informal care is voluntary, unpaid care provided to a person within the informal caregiver’s social network, who has physical, mental or psychological limitations. Informal care is limited to support that goes further than might be considered usual in a personal relationship, that is, care tasks that—in absence of a health problem—would be fulfilled by the person him or herself (eg, household work, personal care) or would not be needed (eg, physical support, nursing care).’

As in most countries, a major share of home care in the Netherlands is provided by informal caregivers. Recent estimates suggest that more than one in four Dutch citizens aged between 16 and 69 years provides informal care: the vast majority provide support on a weekly basis.²⁸ Although the casemix classification is intended for prospective, per-case payment of formal care, ignoring the impact of informal caregiving could result in unsound decision making.²⁹ Thus, we will also develop dependent variables that include both formal and informal care cost. During each needs (re)assessment within the inclusion period, registered nurses will provide an estimate of how much informal care (in hours) was received by the client in the previous 7 days, based on the definition of informal care provided above. Assuming relatively stable informal care needs per care episode, this estimate is then extrapolated forward to construct two additional dependent variables: mean weekly total cost of (formal and informal) home care during the 4-week and 13-week periods following the needs assessment. In line with previous research, we will set the weight for informal care time to around 0.5 to account for different productivities between informal and formal, and will assess the impact of the exact weighting on the predictive accuracy of the models.^{7 25}

Independent variables (predictors)

Casemix predictors are based on: (1) a CM-SF questionnaire; (2) (nursing) classification data and (3) client demographics and care category.

CM-SF questionnaire

A recent survey among Dutch home care nurses found that they consider physical functioning, cognitive functioning and illness prognosis (in particular, whether a client is terminally ill) as the most relevant predictors of home care use.²⁰ Similarly, the New Zealand HCSS casemix classification emphasises the importance of physical and cognitive functioning, self-reliance (in terms of Instrumental Activities of Daily Living (IADL)) and informal caregiver burden.⁴ Since these characteristics are not routinely and/or uniformly assessed across all (nursing) classification systems, we developed a CM-SF questionnaire to enable their possible inclusion for the purposes of casemix classification.

The development and psychometric assessment of the CM-SF will be described in detail in a separate paper. In short, a preliminary version was created by drawing

items from existing, validated questionnaires, which measure the predictors described above. If a validated questionnaire was not available for a specific predictor of interest—or was considered too elaborate or complex to be incorporated into the CM-SF—a tailor-made item was added. The preliminary version of the CM-SF was tested and revised based on four focus groups with three to five registered nurses each, as well as a small-scale pilot involving approximately 20–25 clients. The final version is included in online supplementary file 1. It comprises 11 casemix items in five categories: (1) illness prognosis; (2) functional status, based on the Katz Index of Independence in ADL³⁰; (3) self-reliance, based on two measures from the Lawton IADL scale³¹; (4) cognitive skills for daily decision making, based on one item from the InterRAI Home Care Assessment and (5) informal care burden. The CM-SF will be completed by a registered nurse directly after each needs (re)assessment of a new or existing client (ie, at the start of a care episode).

(Nursing) classification data

Providers participating in this study use the two most common (nursing) classification systems in Dutch home care, that is, the Omaha system (n=2) and NANDA-I system (n=2) for client needs assessment. Omaha comprises a Problem Classification Scheme, which consists of signs and symptoms grouped hierarchically into 42 problem classes across four domains: environmental, psychosocial, physiological and health-related behaviours.³² NANDA-I organises nursing diagnoses into different categories, with three levels: 13 domains, 47 classes and 216 diagnoses. For every nursing diagnosis, there are defined subjective or objective characteristics.³³

For all clients included, we will use (nursing) classification data produced by a registered nurse at the start of each care episode: specifically, the nursing diagnoses, defining characteristics and related factors of NANDA-I, and the problem classes, and signs and symptoms of Omaha. We will distinguish between the NANDA-I diagnoses and Omaha problem classes included in a client’s care plan (which typically have interventions associated with them) from the more extensive and detailed NANDA-I and Omaha assessments recorded for each client.

Demographics and care categories

In terms of demographics, we will use clients’ age (in years), sex (male/female) and four-digit postal area code. Based on the latter, the mean income for the relevant four-digit postal area code will be linked to each client using publicly available data from Statistics Netherlands. Each client’s ‘care category’ is also used as predictor. Since 2017, as part of home care needs (re)assessment, registered nurses are required by health insurers to allocate clients to one of seven possible care categories, on the basis of which category best reflects the expected nature of care provision. Examples include short-term care for frail elderly and chronically ill (<3 months), care

for terminally ill patients, and care for children. The decision tree for allocation to a single care category can be found in a recent study by De Korte *et al.*¹⁸

Sample size

This is an exploratory study, and therefore, the numbers of participants are chosen on pragmatic grounds. We will include as many eligible clients as possible from the participating home care providers within the inclusion period. However, to get a provisional estimate of the number that is reasonably needed, a comparison was made with the recently developed HCSS casemix classification from New Zealand,⁴ which also uses (nursing) classification data. The HCSS algorithm for complex clients is the most relevant to Dutch home care, since these clients receive personal and nursing care as covered by Dutch health insurance. It distinguishes eight clusters. We used simulations (see online supplementary file 2) for our sample size calculations. Our analysis suggests that a minimum sample size of at least 1500 clients is needed, although more would be better still. Given providers' client numbers, we expect to meet the minimum sample size by the end of the inclusion period.

Data collection, processing and cleaning (data management)

All data necessary to operationalise the dependent and independent variables of interest are collected from providers' EHRs and administrative databases. For this purpose, a Minimal Data Set (MDS) is defined, specifying which data fields need to be extracted. The MDS is available in online supplementary file 3. A specialist EHR consultant will make sure that the MDS—including data from the CM-SF, which has been integrated into providers' EHR—is extracted correctly and uniformly. All data will be anonymised, with only a unique client identifier for data merging purposes. The data are transferred to the Dutch Healthcare Authority using a secure file transfer portal, and then processed and analysed using the statistical software package R.

Missing value analysis is used to quantify missing data (ie, clients with a care plan, but without CM-SF or vice versa) and to identify the reason for missing data. Also, outliers in the dependent variables will be identified and

checked in the raw data for anomalies. For the analyses comparing the various sets of predictors, complete cases are included; incomplete cases are checked to assess whether missing data are missing at random.

Statistical analysis methods

Descriptive statistics (eg, frequencies, means and SD, and medians and IQRs, as appropriate) are used to describe various dependent and independent variables across episodes of care.

For model building, datasets are created that consist of a single dependent variable combined with different sets of predictors. Table 2 lists all planned combinations. Previous work on similar data showed that there are clients with either a comparatively low care use (defined as <4 hours during a care episode) or high care use (defined as >40 hours per 4 weeks care episode or >120 hours per 13 weeks episode).¹⁸ It is not uncommon to exclude such clients from a prospective payment system based on casemix classification and instead to reimburse their care based on, for example, FFS.³⁴ To enable such policy choices, models will be developed on the full dataset and on a subset of data excluding these cases.

Each dataset is processed through five modelling pathways, shown in figure 1. Each modelling pathway consists of three stages, that is, variable preprocessing and selection, model building and performance evaluation.

Variable preprocessing and selection

Variable preprocessing consists of 'one-hot' encoding of some or all items in the CM-SF. Seven CM-SF items contain ordered categorical responses; four items have non-ordered categorical responses.

NANDA-I contains over 2000 defining characteristics. Including all these variables as predictors is not desirable because some characteristics are very rare; moreover, considerable practice variation in needs assessment is expected at this level of detail. However, a rare characteristic may still be valuable if it has a significant effect on the outcome. Based on previous work,¹⁸ we expect a few hundred predictors to be left over if we initially select all NANDA-I or Omaha variables present in at least 2% of all clients. This also has the practical benefit of reducing

Table 2 Planned combinations of predictor sets

	All providers (n=4)	Providers using Omaha (n=2)	Providers using NANDA-I (n=2)
CM-SF items	+	+	+
CM-SF items, demographics and care categories	+	+	+
CM-SF items, demographics, care categories and care plan-specific classification data*	–	+	+
CM-SF items, demographics, care categories and complete classification data†	–	+	+

*Omaha problem classes and NANDA-I diagnoses.

†Omaha problem classes, and signs and symptoms, and NANDA-I diagnoses, defining characteristics, and related factors. CM-SF, Casemix Short-Form.

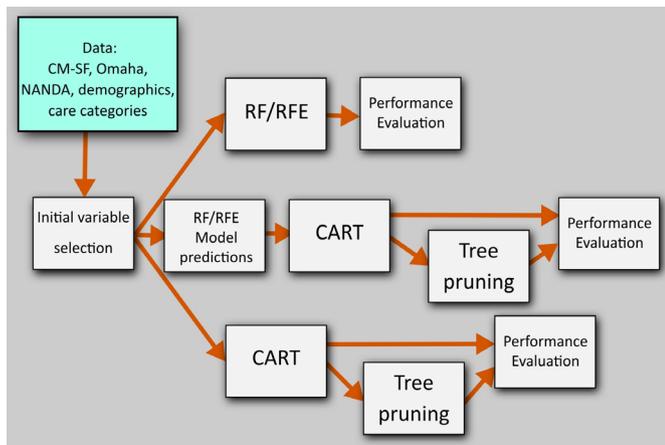


Figure 1 Schematic representation of the model building and evaluation process. CART, classification and regression trees; CM-SF, Casemix Short-Form; RF/RFE, random forest-recursive feature elimination.

the computational cost of the downstream modelling steps. Assuming a few thousand observations, this translates into around 10 data points per binary predictor. For predictions of a binary outcome, the literature recommends using at least 10–20 events per variable³⁵; for a continuous regression outcome with binary predictors, no such rule of thumb was found. Having only 10 data points per predictor may be on the low side, given the flexible nature of tree-based methods.³⁶ However, this will only affect models using complete classification data (see table 2): the potential amount of overfitting is quantified by building a model from the NANDA-I/Omaha data of a single provider, and predicting on the data of another provider.

Model building

As part of the modelling pathways, we will complete the following steps.

1. Fit a random forest model, optimised by first performing recursive feature elimination, followed by tuning the algorithm parameters (mtry, nodesize). This will give us an estimate of the maximum predictive value contained in the set of predictors considered.
2. Fit a random forest model, use this model to generate a prediction for each care episode and then fit a classification and regression trees (CART) model on these predictions. This approach is termed ‘single tree approximation’ (STA) for a black box machine learning model.³⁷ It is expected to result in an interpretable model, for example, a single decision tree, which suffers less from the known drawbacks associated with trees derived directly from training data (eg, sensitivity to small differences in input data).³⁸
3. Fit a CART model directly on the data, with relevant features selected during variable selection. This is arguably standard practice for casemix development and serves as baseline reference to compare with the Random Forest approach in step (1) and the STA approach in step (2).

For steps (2) and (3), we will then either: (1) prune, using cost complexity pruning (we will follow the one-standard-error rule variant described in detail in Kuhn and Johnson³⁸), the CART model for optimal prediction accuracy; or (2) prune the CART model to a prespecified number of terminal leaves (20, 50 and 150), resulting in a prespecified number of casemix clusters. This allows us to quantify the loss of accuracy when model complexity is reduced below the level of complexity required for optimal out-of-sample prediction accuracy.

The first two model building steps include additional variable selection in the form of random forest-based recursive feature elimination.³⁸ This is a stepwise backward procedure that removes variables based on the importance of the permutation variable, and stops as soon as the predictive accuracy starts to decrease by more than one SE. Permutation variable importance is used instead of the default Gini importance to avoid bias towards continuous variables and variables with many categories.³⁹

Performance evaluation

Model performance is assessed using four different statistical performance measures, that is, R-squared (R^2), Cumming’s prediction measure (CPM), mean average prediction error (MAPE) and root-mean squared error (RMSE). R^2 and RMSE are based on a squared error loss function, and quantify goodness-of-fit attributing a more than proportional (ie, quadratic) higher loss value for predictions that are far off (‘outliers’) compared with predictions close to the actual outcome. CPM, a performance measure that is increasingly popular for evaluating risk adjustment models, and MAPE quantify goodness-of-fit when the loss function is simply proportional to the distance between prediction and actual outcome.^{40 41} All performance measures are reported to facilitate policy discussions. Performance measures are evaluated using cross-validation on data not used to build the casemix classification, so as to prevent overfitting. In addition, relative cost homogeneity within casemix groups is quantified using the coefficient of variation. For all models, diagnostic plots will be made that split the prediction error by decile of the values of the dependent variable.

Model validation

Internal validation is addressed by using cross-validation at various stages of the modelling pathways. To provide some indication of external validity, we will take advantage of the fact that we have four participating providers in four separate regions. This allows us to fit models to three providers and test them on the fourth provider, and repeat this process until every provider has been left out of model building once. For models containing Omaha or NANDA-I predictors, this comes down to fitting the model on one provider and testing on the other provider. This procedure has been termed ‘internal–external validation.’⁴² External validation is not part of this study.

Patient and public involvement

No patient involved.

Ethics and dissemination

Data are processed without the explicit consent of participants on the basis of the legal obligation of the Dutch Healthcare Authority to supervise healthcare markets (article 16(a), Healthcare Market Regulation Act). Any information that could be used to identify individual persons is removed or anonymised prior to data collection. Participating providers have been provided with materials to inform clients about the purpose of this study in accordance with the EU general data protection regulation. An opt-out form was created to allow nurses to flag clients who wish to opt out of the study in the EHR.

The initial research findings are expected in 2020 and will serve as input for the development of a new payment system for home care in the Netherlands, to be implemented at the discretion of the Dutch Ministry of Health, Welfare and Sports. Results will also be published in peer-reviewed publications and presented at national and international conferences.

Twitter Arianne Mathilda Josephus Elissen @ArianneElissen

Acknowledgements The authors thank the four homecare providers for participating in this study, and in particular the nurses for administering the CM-SF questionnaire. Also, the authors thank Mylette Lanting (senior consultant at Quattri) for her help with designing the minimal data set specification and subsequent data extraction from the databases of the participating providers. Finally, the authors wish to thank the academic partners from Utrecht University and the HU University of Applied Sciences Utrecht (Nienke Bleijenberg and Jessica Veldhuizen) for their valuable input on this work during meetings of the Dutch Healthcare Authority's Scientific Programme on Home Care Nursing.

Contributors AMJE and GSV drafted the study protocol with input from all authors. GSV and MHdK provided statistical advice and support. GSV, MHdK and LCvdW contributed to the acquisition and interpretation of data for the work. JS conceived the study. AOEvdB, AMJE and SM developed the CM-SF questionnaire with input from all authors. MCM, DR and JS were in charge of overall direction and planning. All authors provided critical feedback, helped revise the manuscript and are accountable for the accuracy and integrity of all aspects of the work.

Funding This work is supported by research grants from the Dutch Healthcare Authority and from home-care organisation MeanderGroep Zuid-Limburg, the Netherlands.

Disclaimer The views expressed in this article are those of the authors and not necessarily those of the Dutch Healthcare Authority (NZa) or MeanderGroep Zuid-Limburg, the Netherlands.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval An accredited Medical Research Ethics Committee reviewed this study and classified it as research not subject to the Dutch Medical Research Involving Human Subjects Act (reference number 2019–1144).

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Arianne Mathilda Josephus Elissen <http://orcid.org/0000-0001-9795-8095>

REFERENCES

- Costa AP, Poss JW, McKillop I. Contemplating case mix: a primer on case mix classification and management. *Health Manage Forum* 2015;28:12–15.
- Jegers M, Kesteloot K, De Graeve D, et al. A typology for provider payment systems in health care. *Health Policy* 2002;60:255–73.
- Langenbrunner JC, Cashin C, O'Dougherty S. *Designing and implementing health care provider payment systems*. Washington, DC: The World Bank, 2009.
- Parsons M, Rouse P, Sajtos L, et al. Developing and utilising a new funding model for home-care services in New Zealand. *Health Soc Care Community* 2018;26:345–55.
- Lave JR. Developing a Medicare prospective payment system for inpatient psychiatric care. *Health Aff* 2003;22:97–109.
- Kobel CTJ, Bellanger M, Pfeiffer KP. DRG systems and similar patient classification systems in Europe. In: Busse R, Geissler A, Quentin W, eds. *Diagnosis-Related groups in Europe: moving towards transparency, efficiency and quality in hospitals*. Maidenhead: Open University Press, 2011: 37–58.
- Poss JW, Hirdes JP, Fries BE, et al. Validation of resource utilization groups version III for home care (RUG-III/HC): evidence from a Canadian home care jurisdiction. *Med Care* 2008;46:380–7.
- Ashcraft ML, Fries BE, Nerenz DR, et al. A psychiatric patient classification system. An alternative to diagnosis-related groups. *Med Care* 1989;27:543–57.
- Starfield B, Weiner J, Mumford L, et al. Ambulatory care groups: a categorization of diagnoses for research and management. *Health Serv Res* 1991;26:53–74.
- Manton KG, Hausner T. A multidimensional approach to case mix for home health services. *Health Care Financ Rev* 1987;8:37–54.
- Masso M, Allingham SF, Banfield M, et al. Palliative care phase: inter-rater reliability and acceptability in a national study. *Palliat Med* 2015;29:22–30.
- Slobbe LCJ, Wong A, Verheij RA, et al. Determinants of first-time utilization of long-term care services in the Netherlands: an observational record linkage study. *BMC Health Serv Res* 2017;17:626.
- Lee Y, Penning M. The determinants of informal, formal, and mixed in-home care in the Canadian context. *J Aging Health* 2019;31:1692–714.
- Kadushin G. Home health care utilization: a review of the research for social work. *Health Soc Work* 2004;29:219–44.
- MedPac. *Home health care services payment system*. Washington DC: MedPac, 2017.
- Teenier P. 2008 refinements to the Medicare home health prospective payment system. *Home Healthc Nurse* 2008;26:181–4.
- James ML FB, Goodell M, Wellens N. *Design principles for HCBS case mix: a primer*. Ann Arbor, MI: University of Michigan, Institute of Gerontology, 2015.
- De Korte MH, Verhoeven GS, Elissen AMJ, et al. *Using machine learning to assess the predictive potential of standardized nursing data for home healthcare case-mix classification*, forthcoming.
- van Dorst J, Rosendal H, Metzelthin SF. Maakt Het uit WAT JE gebruikt? *Nederlands Tijdschrift voor Evidence Based Practice* 2017;15:17–19.
- van den Bulck AOE, Metzelthin SF, Elissen AMJ, et al. Which client characteristics predict home-care needs? results of a survey study among Dutch home-care nurses. *Health Soc Care Community* 2019;27:93–104.
- Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
- Maurits EEM. *Autonomy of nursing staff and the attractiveness of working in home care*. University of Utrecht: Utrecht, 2019.
- Vektis Intelligence. *Wijkverpleging. Zorgprisma Publiek*. The Netherlands: Vektis Intelligence, 2018.
- Kroneman MBW, Van den Berg M, Groenewegen P, et al. *Health systems in transition: the Netherlands*. Brussels, Belgium: European Observatory on Health Systems and Policies, 2016.
- Björkgren MA, Fries BE, Shugarman LR. A RUG-III case-mix system for home care. *Canadian Journal on Aging / La Revue canadienne du vieillissement* 2000;19:106–25.
- PricewaterhouseCoopers. *Onderzoek tariefherijking verpleging en verzorging*. Amsterdam: PricewaterhouseCoopers Advisory NV, 2017.
- Metzelthin SF, Verbakel E, Veenstra MY, et al. Positive and negative outcomes of informal caregiving at home and in institutionalised long-term care: a cross-sectional study. *BMC Geriatr* 2017;17:232.
- De Boer API, De Klerk M. *Werk en mantelzorg: kwaliteit van leven en Het gebruik van ondersteuning OP Het werk*. Den Haag: Sociaal en Cultureel Planbureau, 2019.



- 29 Hoefman RJ, van Exel J, Brouwer WBF. The monetary value of informal care: obtaining pure time Valuations using a discrete choice experiment. *Pharmacoeconomics* 2019;37:531–40.
- 30 Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc* 1983;31:721–7.
- 31 Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179–86.
- 32 OMAHA. The OMAHA system. solving the clinical data-information puzzle, 2018. Available: <http://www.omahasystem.org/overview.html>
- 33 Herdman TH. *NANDA International, Nursing Diagnoses: Definitions & Classification 2012-2014*. West Sussex, UK: John Wiley & Sons Ltd, 2012.
- 34 Centers for Medicare and Medicaid Services. *Home health prospective payment system (HH PPS) rate update for Calendar year (cy), 2019* MLN Matters Number: MM10992. [Available from: <https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/MM10992.pdf>
- 35 Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373–9.
- 36 van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC Med Res Methodol* 2014;14:137.
- 37 Guidotti R, Monreale A, Ruggieri S, et al. A survey of methods for explaining black box models. *ACM Comput Surv* 2019;51:1–42.
- 38 Kuhn M, Johnson K. *Applied predictive modeling*. Heidelberg: Springer Nature, 2013.
- 39 Strobl C, Boulesteix A-L, Zeileis A, et al. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 2007;8:25.
- 40 Cumming RB, Cameron BA. *A comparative analysis of claims-based methods of health risk assessment for commercial populations*. Minneapolis, MN: Milliman US, Inc, 2002.
- 41 van Veen SHCM, van Kleef RC, van de Ven WPMM, et al. Is there one measure-of-fit that fits all? A taxonomy and review of measures-of-fit for risk-equalization models. *Med Care Res Rev* 2015;72:220–43.
- 42 Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 2016;69:245–7.