Assessment of variation in long-term outcomes of integrated care initiatives in Dutch health care

Katalin Katona, Malou Dorine Menting and Ylva Michelle Pisters

Abstract

Abstract:

Introduction: The care for many patients with diabetes mellitus type 2 in the Netherlands, is contracted by a local care group. The healthcare providers, who collectively shape a care group, provide protocolled diabetes care. Differences exist between care groups in terms of their organizational and financial arrangements. These differences may result in variation in outcomes. The aim of this study is to assess whether variation in healthcare costs, diabetes complications and related hospital admissions on the level of care groups exist.

Methods: A quantitative cohort study was conducted. Patients who used diabetes medication (more than 180 days of defined daily doses per year) for the first time between the years 2014 and 2019 were included. Data were extracted from health insurance claims between 2014 and 2019. Generalized linear mixed models were used to analyse patient variation in healthcare costs (two and six years follow-up), diabetes-related complications and hospital admission days. Intraclass correlation coefficients were calculated to estimate the amount of variation that was attributable to the care groups.

Results: A large variation in outcome variables was observed between patients and a small variation between care groups. The intraclass correlation coefficient for long-term costs was 0.4%; for short-term costs between 0.1% and 0.3%; for complications 1% and for hospital days 4%.

Discussion: A large variation between patients with diabetes mellitus type 2 exists in terms of their healthcare costs and complications. In our study, care groups accounted minimally for this variation. A generalized linear mixed model in combination with year cohorts is a tool to study variations in the long-term outcomes of integrated care initiatives.

Keywords

Diabetes mellitus type 2, integrated health care systems, health care costs, complications of diabetes mellitus, the Netherlands, patient care bundles/economics

Introduction

The steady increase of patients with chronic diseases is putting pressure on the healthcare system and healthcare budgets. Diabetes mellitus (DM) is one of those highly prevalent chronic diseases. Ninety percent of patients with DM have DM type 2 (DM2). This group is at high risk of diabetes complications. The treatment of these diabetes complications accounts for the majority of healthcare spending for patients with DM2. These diabetes complications seem to be largely preventable through effective diabetes management.

As a potential solution to the constantly rising costs of treating DM2, integrated care groups (hereafter: care groups) and bundled payments (BP) for chronic diseases were introduced in the Netherlands in 2010. The care for nearly all patients with DM2 treated in the primary care in the Netherlands is contracted by a care group. The healthcare providers that collectively shape a care group, provide highly protocolled diabetes care. The different healthcare providers that can form a care group together

Dutch Healthcare Authority, Utrecht, The Netherlands
Department of Health Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Corresponding author:
Malou Dorine Menting, Dutch Healthcare Authority, Utrecht, 3502 GA, The Netherlands.
Email: mmenting@nza.nl
are diverse: general practitioners (GP), nurse practitioners, internists, specialist nurses, podiatrists, dietitians, ophthalmologists, physiotherapists, psychologists, pharmacists, and medical podiatrists. Which healthcare providers form together the integrated care group differs between care groups, but all care groups have a large representation of GPs: the mean number of GPs per care group in 2013 was 77. Most care groups are therefore run by local GP associations. GP associations are also generally the main contractors.

In the Netherlands, health care insurers sign contracts with health care providers. To encourage the integration of different health care providers within the care group, health care insurers contract and purchase the different care components of the DM disease management program as a single package from the care group. This payment is for the full range of services for the patient. Care groups, in turn, contract general practitioners and other health care providers like dieticians and specialists. This is known as a bundled payment (or BP) and replaces the traditional Fee For Service (FFS) approach in which the different care components are contracted by the individual health care providers.

A care group is thus a provider-led organization that is clinically and financially responsible for all patients in a disease-specific program. The assumption of integrated care is that the quality of care is improved through better coordination within the care group. This ideally leads to better patient health and lower healthcare costs. The Dutch care groups are comparable to clinical commission groups in England and Accountable Care Organizations in the United States.

Indeed, previous research based on studies from Europe, Australia and Asia showed that integrated care seems to decrease health care costs or medical spending growth and improves health outcomes. Also, a different payment scheme than FFS seems a powerful tool to stimulate integrated care and influence health care expenditure growth. However, previous research is inconclusive about the effect of bundled payments for integrated care provided by care groups, in terms of patients’ clinical outcomes and healthcare costs in the Netherlands. A few years after the introduction of care groups with BP in the Netherlands, small improvements were detected in outcome measures such as systolic blood pressure and cholesterol and improved coordination of care, but health care expenditure seems to have increased, especially for multimorbid patients.

The content of diabetes care is based on a national standard. It is based on recent and evidence-based insights for coordinated, individual prevention and care. The standard is functional, so it only describes what care needs to be provided, not by whom. However, differences exist between care groups in terms of their organizational and financial arrangements. These differences may result in an undesirable variation in outcomes. Every year, care groups in the Netherlands are benchmarked (on a voluntary basis) regarding aspects of the care they provide. These benchmarks show that the care is highly protocolized and the variation between care groups has decreased over the years. Nevertheless, there is still some undesirable variation between care groups on certain indicators such as the percentage of patients with a low LDL-cholesterol level, lipid-lowering medication and foot and fundus examinations. Quality agreements between various health insurers and care groups also differ, leading to variation between care groups.

Variation is not per definition undesirable. The current transition towards personalized, patient centred, and value-based care increases differences between care groups in the care they provide. Studying patient outcome variation between care groups helps to redesign diabetes care towards a more patient centred, integrated approach. Multilevel modelling is an important and often used statistical technique to gain information from the variance of outcome or cost variables between providers and so find ways to improve the disease management program. The key idea is to isolate the variation in outcome variables due to justifiable factors, such as differences in the patient case mix. If there is still a considerable variation attributable to providers, the practices of the best performing providers can provide information on the best way to achieve quality and cost improvements in diabetes care. The research that is most comparable to the present study, investigates spending variation between integrated care organizations with a multilevel model (GLMM). This study showed a small, but significant variation in patients’ healthcare costs between care groups. It used data from one year and clinical patient outcomes were not studied. Given the clinical course of DM2 (many DM2-related complications develop later), it is important to track the patients over a longer period of time.

The present study assessed whether variations exist in long-term (6 years) healthcare costs, diabetes-related complications and hospital admissions of patients with DM2 between care groups in the Netherlands. Additionally, the potential variation over time was examined, using short-term healthcare costs (2 years) of the patients. This study is the first attempt to assess variation in the long-term outcomes of integrated care initiatives in Dutch health care.

Methods

Study design and data collection

The design used in this study is a quantitative cohort study. The individual-level dataset included healthcare claims, patient characteristics (collected by Vektis, the information center established by the Dutch insurers), social-economic status (SES) and pharmaceutical cost groups (FCGs)
(collected for the Dutch risk equalization system) for the years 2012–2019. This dataset was provided by the Dutch Healthcare Authority (Nederlandse Zorgautoriteit, NZa). The dataset was pseudonymized. Public data from CBS (Statistics Netherlands) was used to account for the degree of urbanization in the patients’ region.

**Study population**

Patients with DM2 were selected if they received integrated care between 2014 and 2019. Healthcare declaration codes were used to identify which patients received integrated care by a care group (codes are available in the Supplemental Material). Further, patients were selected who started to use more than 180 defined daily doses (DDD) of DM medicines per year for the first time between 2014 and 2019. This threshold for medication use coincides with the threshold used in the Dutch risk equalization system. Patients above this threshold generally have higher healthcare costs than patients without DM.21 Data from 2012 to 2013 was used to exclude patients whose medication usage already exceeded this threshold before 2014. The patients included were classified into cohorts 2014–2019, with the year indicating the first year in which their medication usage exceeded the threshold. All analyses were based on cohort 2014, except for the analyses on short-term healthcare costs.

Patients with DM type 1 (DM1), patients with gestational diabetes and patients younger than 18 years were not eligible for care by a care group and thus expected to be automatically excluded from this study. Patients younger than 18 years were explicitly excluded. The following patients were also explicitly excluded: patients treated by multiple care groups during the study period or patients treated by a care group that did not exist continuously throughout the study period.

**Variables**

The following outcome variables were constructed: total healthcare costs, number of diabetes-related complications, and total days of hospital admissions related to complications. Total healthcare costs were the total costs per patient in the following sectors: general practitioner (including first-line multidisciplinary care), pharmaceutical care, hospital care, medical aids, podiatry, and dietetics.

For the main analysis of long-term costs, the total 6 years (2014–2019) cost was summed up into a single outcome measure. The additional analysis of short-term costs used the sum of total costs in the first 2 years of the cohort as an outcome variable.

An a-priori list of 13 possible diabetes-related complications was created based on literature.1–13 Diagnosis treatment combinations (DTCs) within the relevant target groups, that is, specialism-diagnosis combinations, corresponding to these complications were selected. DTCs that only included screening or diagnostics for the complication were not included (list complications, target groups and DTCs available in the Supplemental Material). This strategy was intended to capture serious complications that needed medical care and thus involved high hospital costs. Every complication was counted once per patient. The total number of diabetes-related complications was summed for the time period 2014–2019.

Hospital admission days were extracted based on the care activity codes (included in the Supplemental Material) within the complications selected. Activity codes were only available from 2016 onwards. Thus hospital admission days in the period 2016–2019 were included.

**Confounding variables**

Based on previous research and literature on the factors that are correlated with diabetes-related complications and healthcare costs, the following variables we controlled for: gender, age, socioeconomic status (SES), degree of urbanization (as a proxy for capacity and competition of hospitals), and the number of pre-existing comorbidities.23–28 These characteristics are known to influence healthcare costs, the likelihood of complications and hospitalizations. Socioeconomic status was defined as in the Dutch risk equalization system in 2014. Pre-existing co-morbidities were counted based on pharmaceutical cost groups (FCGs) for the patient in the year of inclusion in the study. FCGs are defined as in the risk equalization system and are based on medication use for different diseases. Supplemental Appendix 5 describes all the FCGs that were included. The death of a patient during the study period was taken into account by including a dummy variable in the analyses of total costs, and by including the time of exposure (time at-risk) in the analysis of complications and hospital admissions.

At the care group level, the size of the care group was controlled for expressed in the total number of diabetes patients, and the primary health insurer. The primary health insurer was used as a proxy for the content of the care group’s protocols and the disciplines participating in the group. The care groups’ primary healthcare insurers were grouped into six groups. There were four groups for each of the four biggest healthcare insurers; one group was a combination of two healthcare insurers (both primary healthcare insurers for the same care groups) and one group included three smaller healthcare insurers.

**Analyses**

The data had a nested structure because patients were nested within care groups. Because of this nested structure and the non-normal distribution of the outcome measures, a Generalized Linear Mixed Model (GLMM) was used.29
The levels in the model were the patients and the care groups. In the analysis of total costs, a log-normal distribution with identity link function was used, while in the analysis of the number of complications and admission days, a negative binomial distribution with log link function was assumed. In the latter case, the dispersion parameter of the model was assessed to check whether the negative binomial distribution fitted the data better than a Poisson distribution.

The total costs (long-term and short-term) were estimated using the following model:

\[ Y_{ij} \sim \text{Lognormal}(\mu_{ij}, \sigma^2) \]

\[ \mu_{ij} = \beta X_{ij} + \gamma Z_i + \alpha_j \]

\[ \alpha_j \sim N(0, \tau^2) \]

To estimate the number of complications and admission days, the following model was used:

\[ Y_{ij} \sim \text{NB}(\lambda_{ij}, \phi) \]

\[ \log(\lambda_{ij}) = \beta X_{ij} + \gamma Z_i + \alpha_j + \log(T_{ij}) \]

\[ \alpha_j \sim N(0, \tau^2) \]

\( Y_{ij} \) is the outcome measure; total costs, the number of complications or the total days of hospital admissions, respectively, of patient \( i \) nested in care group \( j \). \( Y_{ij} \) follows a lognormal distribution in the case of costs and a negative binomial distribution in the case of the other two outcome measures, conditional on the covariates and the care-group-specific random effect. \( \sigma^2 \) is the variance of the lognormal distribution, while \( \Phi \) is the dispersion parameter of the negative binomial distribution. \( X_{ij} \) is the vector of the covariates gender, age, SES, degree of urbanization, number of pre-existing comorbidities and deaths during the study period in case of the analysis of total costs, of patients \( i \) nested in care groups \( j \). This also includes one constant intercept term. \( \beta \) is the vector of regression coefficients of the patient covariates. \( T_{ij} \) is an offset term that indicates the time at-risk for each patient in the case of the analyses of complications and the number of hospital days. The coefficient of an offset term is always 1, so is not presented as a coefficient in the final model. \( Z_i \) is the vector of covariates size of care group and primary health insurer of care groups \( j \). \( \gamma \) is the vector of regression coefficients of the care group covariates. The fixed effects are the patient covariates and the care group covariates. The random effect is the intercept of the care group level, \( \alpha_j \), which is distributed normally with mean 0 and standard deviation \( \tau \). The random effect was the primary focus of this study, quantifying the variation among care groups. The statistical software package R, version 4.0.2, was used. Data cleaning, manipulation, and descriptive statistics were mainly carried out using the package ‘data.table’. Regression analyses were based on Bayesian inference, using the ‘brms’ package. The methods and recommendations for Bayesian inference by Gelman and van der Schoot were used. Multilevel modelling in the ‘brms’ package was based on Bürkner and Nalborczyk. The priors were set on the defaults for the package, which were weakly informative. This assumption was checked visually and by carrying out a sensitivity analysis.

First, a crude model was run without the random intercept for the care group. Then the model was extended with the patient confounding variables, and lastly with the care group confounding variables. The variation of outcome measures that can be attributed to the care group level, was based on the Intraclass Correlation Coefficient (ICC). The ICC was calculated by dividing the between-unit variance by the total variance for Gaussian models. However, in the negative binomial mixed models, the between-unit variance was absent because of the different scales of the cluster error variance and the outcome measures. Variance Decomposition (VD) is the recommended alternative in this case. For the VD, draws from the posterior predictive distribution with and without random effects were calculated first. Then the variances for the draws were calculated. Lastly, the VD was the ratio between the difference of the variances and the variance of the draws from the distribution with random effects. As a verification, the ICC was calculated with a different, rough method proposed by Nakagawa et al. using 4000 posterior samples for lambda (i.e. expected number of complications) 1 and 2.

Additionally, a test was run on whether the model with the care-group-specific random intercept was preferable to a model without a random intercept. This was done by comparing the \( R^2 \) of both models, and by performing a leave-one-out cross-validation (LOO) comparison. LOO is a method for testing Bayesian models that estimate the predictive accuracy of the model using the log-likelihood of posterior simulations. The outcome is the expected log predictive density (ELPD). In the LOO compare output, the difference in expected predictive accuracies between the models is expressed by the ELPD difference.

The variables of patient age and care group size were standardized to help the model-fitting process. For all models, chain convergence was reached and checked through visual observations of the trace plots, \( \hat{\gamma} \) and substantial effective sample sizes (ESS). The distribution of the predicted versus observed values was plotted to check the model fit. The appropriateness of the default priors was verified by sampling and plotting the priors and comparing them with the posteriors of the same parameters, and by sensitivity analyses. Lastly, it was verified that the negative binomial distribution was preferable to the Poisson distribution in the models on complications and admissions.
hospital admissions since the 95% credible interval of the dispersion parameters did not include zero.

Sensitivity analyses

Three sensitivity analyses were conducted: using different priors; running all models on data from cohort 2015; and including the GP practices within the care groups as an extra level in the model.

Results

Patient and care group characteristics

A total of 235,080 patients who used more than 180 defined daily doses of diabetes medication for the first time in the period 2014–2019 were identified. A total of 143,283 patients were included in the study (see flowchart in Supplemental Files). Table 1 presents the descriptive statistics of patient characteristics in the cohorts 2014–2018.

| Cohort     | 2014     | 2015     | 2016     | 2017     | 2018     |
|------------|----------|----------|----------|----------|----------|
| Total in cohort | 35,208   | 31,830   | 26,181   | 25,246   | 24,818   |
| Male (N, %)  | 18,973 (54) | 17,708 (56) | 14,559 (56) | 14,414 (57) | 14,256 (57) |
| Age (Mean, SD) | 65.0 (12.7) | 64.4 (12.6) | 63.9 (12.7) | 63.6 (12.6) | 63.5 (12.5) |
| Number of deaths (N, %) | 5620 (16) | 3814 (12) | 2339 (9) | 1553 (6) | 818 (3) |
| Socio-economic status (N, %)* | | | | | |
| Low         | 8214 (23) | 7273 (23) | 6117 (23) | 5970 (24) | 5864 (24) |
| Medium      | 17,213 (49) | 15,179 (48) | 12,287 (47) | 11,541 (46) | 11,162 (45) |
| High        | 9428 (27) | 8909 (28) | 7102 (27) | 6871 (27) | 6898 (28) |
| Other or missing | 353 (1) | 469 (2) | 675 (3) | 864 (3) | 894 (4) |
| Urbanicity (N, %)# | | | | | |
| Extremely urbanized | 6586 (19) | 5864 (18) | 4610 (18) | 4469 (18) | 4393 (18) |
| Heavily urbanized   | 8842 (25) | 7741 (24) | 6501 (25) | 6257 (25) | 6148 (25) |
| Moderately urbanized | 6924 (20) | 6323 (20) | 5284 (20) | 5113 (20) | 4899 (20) |
| Hardly urbanized     | 6751 (19) | 5984 (19) | 5102 (20) | 4902 (19) | 4848 (20) |
| Not urbanized         | 6019 (17) | 5827 (18) | 4625 (18) | 4426 (18) | 4439 (18) |
| Missing               | 86 (0)   | 91 (0)   | 59 (0)   | 79 (0)   | 91 (0)   |
| Pre-existing comorbidities (N, %) | | | | | |
| 0                      | 15,530 (44) | 13,087 (41) | 10,109 (39) | 9869 (39) | 9966 (40) |
| 1                      | 14,536 (41) | 13,771 (43) | 11,490 (44) | 10,886 (43) | 10,641 (43) |
| 2                      | 4306 (12)  | 4095 (13)  | 3749 (14)  | 3638 (14)  | 3442 (14)  |
| 3                      | 743 (2)    | 776 (2)    | 734 (3)    | 724 (3)    | 666 (3)    |
| 4                      | 87 (1)     | 92 (0)     | 90 (0)     | 113 (0)    | 87 (0)     |
| 5                      | 6 (0)      | 7 (0)      | 8 (0)      | 15 (0)     | 15 (0)     |
| 6                      | 0 (0)      | 2 (0)      | 1 (0)      | 1 (0)      | 1 (0)      |
| Years at risk          | | | | | |
| 1                      | 338 (1)    | 279 (1)    | | | |
| 2                      | 1062 (3)   | 901 (3)    | | | |
| 3                      | 1100 (3)   | 854 (3)    | | | |
| 4                      | 1054 (3)   | 896 (3)    | | | |
| 5                      | 1048 (3)   | 28,900 (91) | | | |
| 6                      | 30,06 (87) | | | | |

*Missing values and other categories than 'low', 'medium', 'high' were excluded from the regression analysis. Socio-economic status as in 2014 was used for all cohorts.

#Missing values were excluded from the regression analysis. Note. The variable 'years at risk' was only used for analyses based on cohorts 2014 and 2015.

Description of outcome variables

Almost half of the patients’ healthcare costs were attributable to the hospital sector (see Supplemental Files). The other half of the healthcare costs were mainly incurred in the pharmaceutical care and general practitioner care
sectors. Low costs were observed in the podiatry and dietetics sectors.

Within the 6-year study period, 36.8% of patients had at least one complication but only 12.7% were hospitalized for at least one day due to those complications. For more details see Supplemental Files.

Regression results in long-term costs

Table 2 describes the results of the models on long-term log-transformed healthcare costs. Model 1 (unadjusted model) indicates that the average cost per patient was €25,931 (exp(9.74 + 0.92²/2)) over the 6-year study period. In Model 2, age, sex, death in the study period, socio-economic status, and pre-existing comorbidities explained about 6% of the variation. Care group characteristics (i.e. size and primary insurer of the care group) did not explain additional variance in the patients’ healthcare costs (model 3). In this final model, the ICC was 0.4%.

The R² of a model without random effects but with all other explanatory variables (not shown in the table) was 0.03, which is certainly lower than the R² of the model with random effects (0.06). The ELPD difference, based on the LOO comparison, for the model without random intercept compared to the full model was −822.9 with a standard error of 40.2. This indicates that the expected predictive accuracy of the model with the random intercept is significantly higher than that of the model without the random intercept, showing that the random intercept for care groups did explain a small part of the variation.

Figure 1 shows the integrated-care-group-specific intercept in the full model (model 3). The 95% CIs of the intercepts overlapped zero in all but five care groups.

Regression results in number of complications and hospital admission days

Table 3 presents the parameter estimates for the full model on the number of complications and hospital admission days.

The ICC in the model on complications was 0.01, calculated with the variance decomposition method. The mean of the ICCs calculated with Nakagawa’s method by using 4000 posterior samples for lambda (i.e. expected number of complications) 1 and 2 was 0.009 and 0.011, respectively, which is comparable to the ICC of the variance decomposition method. In the model on hospital admission days, the ICC was 0.04.

The explained variance by the models on complications and hospitals days, that is, the R², were 0.043 and 0.014, respectively, while the R² of the same models without random intercept were 0.040 and 0.012, respectively. The
ELPD difference, based on the LOO comparison, for the model without random intercept compared to the full model, was $-29.1$ (with a standard error of $7.7$) for complications and $-1.5$ (with a standard error of $4.8$) for hospital days. Both the LOO comparison and the $R^2$ indicated a minor comparative advantage for the models with random intercept (although in the model on hospital days the difference was insignificant), which indicates that the intercepts for the care groups differ from each other.

The care-group-specific random effects (model 3) including the 95% CI ranges are displayed in Figure 2(A). The 95% CIs of the intercepts overlapped zero in all, but four care groups on the outcome complications. The variance between the care groups is larger for hospital days than for complications (Figure 2(B)), but all 95% CIs overlapped zero for the outcome of hospital days.

**Short-term costs**

In the analyses with short-term healthcare costs, the ICC of cohort 2014 was slightly higher than in the other cohorts; however, no consistent trend was visible throughout all the years. See Supplemental Files for the full results of the estimations.

**Sensitivity analyses**

All results were very similar to the primary analysis (see Supplemental Files).

**Discussion**

A large degree of variation exists between patients with DM2 in terms of long-term healthcare costs, diabetes-related complications and hospital days. Care groups accounted only minimally for this variation. This suggests that care groups generally perform equally regarding patients’ outcomes and costs. Our results are only generalizable to patients who used medication for their DM2 (a minimum of 180 daily defined doses for at least one year).

Our findings are consistent with a similar, previous study in this field that uses GLMM and short-term outcomes. In that study based on 2009 data, care groups also had a minimal effect on total healthcare costs. A larger care group effect was found in the variation in diabetes-specific hospital costs. In the present study, the cost of diabetes-specific hospital costs was not assessed separately from the total costs. Patients were followed for 6 years, and it was hard to disentangle the cost of diabetes-specific hospital costs from that of non-diabetes-specific hospital costs because good management of diabetes prevents and delays complications in a wide variety of diseases. Therefore the present study took a broader range of costs into account than those relating to diabetes-specific costs alone. However, a closer look at diabetes-related complications and hospital admission days was taken and no considerable care group effect was observed for the outcome measures in this study.

Currently, 90% of patients with DM2 receive care based on recommended protocols and guidelines. Furthermore, the longer a care program or care group exists, the better its processes and patient outcomes can become. The present study focused on a time period of multiple years after the implementation of integrated care, so all care groups may have reached similar levels of quality. This may explain why care groups did not have a large effect on the variation of patient outcomes in our study. Data about the maturity of the provided integrated care was not available.
Nevertheless, care groups did account for some variation in patient outcomes. This study did not explain the variation. Several factors could explain this variation. Firstly, despite the standardization of diabetes care, the guidelines are still flexible in some respects, such as the fulfillment of specific tasks and roles and the frequency of screening for risk factors. This could lead to differences in the provision of care and consequently in healthcare costs and outcomes. For example, foot care and diagnostics vary widely between care groups, and various studies show this is sometimes managed inadequately. The maturity of integrated care provided by the care group and the number and type of care groups are potentially important factors. Secondly, patients’ healthcare behaviour may differ across care groups, and it was not possible to fully capture this in our study. For example, the proportion of patients who smoke differs significantly between care groups, which could contribute to differences in the development of complications and related costs. Thirdly, differences in hospitalizations and consequently healthcare costs could be affected by the characteristics of regional hospital care, such as payment methods, hospital working culture, and traveling distance. 

A new transformation from standardized care to personalized, tailored care has been introduced to increase the quality of integrated care groups. Additionally, current clinical guidelines in the Netherlands are insufficient for meeting the needs of patients with multiple diseases, which results in maltreatment of adverse events. This calls for the need to move from single-disease integrated care to patient-centred care, integrating multiple care programs for different diseases. When moving away from standardized diabetes care, integrated care groups will have an increasingly large influence on outcomes. This may increase variation in patient outcomes. It is of great importance to closely follow this transformation and to study variation in outcomes again afterwards. The current study showed an analytical approach for assessing long-term variation between care groups. GLMM in combination with year cohorts seems to be a valuable tool for assessing variations in selected, measurable long-term outcomes for integrated care initiatives. Therefore, the present study could help researchers to learn from this analytic approach in their attempts to redesign diabetes care towards a more person-centred, integrated approach.

Our study has several strengths. Firstly, this study included a long (six-year) follow-up of patients. Even so, diabetes complications can take many years to arise, and the effect of care groups on these complications may still be outside the time horizon of our study. However, the specific selection of patients in our study together with the six-year follow-up did increase the likelihood of observing complications. Additionally, with respect to the efficiency of integrated care, reductions in hospital admissions were already observed within one year in other studies.

Secondly, patients were identified based on receiving medication for more than 180 days a year for the first time within one year. This study showed an analytical approach for assessing long-term variation between care groups. GLMM in combination with year cohorts seems to be a valuable tool for assessing variations in selected, measurable long-term outcomes for integrated care initiatives. Therefore, the present study could help researchers to learn from this analytic approach in their attempts to redesign diabetes care towards a more person-centred, integrated approach.

### Table 3. Bayesian multilevel regression model for the effect of care groups on the number of complications and hospital days.

|                         | Model 3 – complications Estimates | Model 3 – hospital days Estimates |
|-------------------------|-----------------------------------|-----------------------------------|
| **Fixed part**          |                                   |                                   |
| Intercept               | −2.54                             | −1.65                             |
|                         | (−2.68 to −2.38)                  | (−2.12 to −1.22)                  |
| Gender                  | −0.12                             | −0.29                             |
|                         | (−0.18 to −0.05)                  | (−0.47 to −0.09)                  |
| Age                     | 0.24                              | 0.40                              |
|                         | (0.20–0.28)                       | (0.31–0.50)                       |
| Urbanicity 2            | 0.04                              | −0.23                             |
|                         | (−0.07 to 0.14)                   | (−0.56 to 0.10)                   |
| Urbanicity 3            | 0.02                              | −0.16                             |
|                         | (−0.09 to 0.13)                   | (−0.52 to 0.20)                   |
| Urbanicity 4            | −0.02                             | −0.15                             |
|                         | (−0.14 to 0.09)                   | (−0.52 to 0.20)                   |
| Urbanicity 5            | −0.02                             | −0.34                             |
|                         | (−0.15 to 0.12)                   | (−0.73 to 0.03)                   |
| SES low                 | 0.11                              | 0.37                              |
|                         | (0.03–0.20)                       | (0.11–0.67)                       |
| SES middle              | 0.08                              | 0.37                              |
|                         | (0.00–0.15)                       | (0.13–0.57)                       |
| Pre-existing comorbidities | 0.18                               | 0.21                             |
|                         | (0.14–0.22)                       | (0.07–0.35)                       |
| Size of care group      | 0.00                              | 0.04                              |
|                         | (−0.06 to 0.07)                   | (−0.10 to 0.22)                   |
| Health insur er 2       | 0.01                              | 0.30                              |
|                         | (−0.25 to 0.26)                   | (−0.26 to 0.90)                   |
| Health insur er 3       | 0.04                              | −0.06                             |
|                         | (−0.17 to 0.23)                   | (−0.52 to 0.46)                   |
| Health insur er 4       | 0.05                              | −0.03                             |
|                         | (−0.16 to 0.21)                   | (−0.42 to 0.43)                   |
| Health insur er 5       | 0.00                              | −0.00                             |
|                         | (−0.20 to 0.22)                   | (−0.47 to 0.60)                   |
| Health insur er 6       | −0.01                             | −0.06                             |
|                         | (−0.23 to 0.20)                   | (−0.66 to 0.55)                   |
| Dispersion parameter    | 1.55                              | 0.04                              |
|                         | (1.45–1.67)                       | (0.04–0.04)                       |
| **Random effects**      |                                   |                                   |
| r (intercept)           | 0.09                              | 0.18                              |
|                         | (0.06–0.12)                       | (0.07–0.29)                       |
| ICC^a                   | 0.01                              | 0.04                              |
| N_care groups           | 86                                | 86                                |
| **Observations**        | 34,811                            | 34,811                            |
| Marginal $R^2$          | 0.039/0.043                       | 0.012/0.014                       |
| conditional $R^2$       |                                   |                                   |

*aBased on variance decomposition.*
time, and thus all patients were in approximately the same stage of their disease. Some general practitioners in care groups diagnose and refer patients to the integrated care program earlier than general practitioners in other care groups. Therefore, simply taking all the patients newly included in the integrated care program for DM2 reduces the comparability of care group populations, producing bias in the results. Our selection made the study population more homogeneous, and thus comparisons between the care group populations were fairer as a result. This method can be an example for other studies struggling with different inclusion criteria for an integrated care program between regions. Moreover, with this method, a group of patients of specific interest when analysing possibilities to cost reduction was identified. Rather than looking at all the patients with DM2, since this group of patients incurs higher healthcare costs than the average citizen.

Thirdly, selection bias was minimized since analyses were performed on data for all Dutch citizens who used medication for DM2 and received care from care groups. Fourth, the three sensitivity analyses all led to the same overall results which strengthen the conclusion.

The study also has limitations. With respect to the medication criterion and the mean age of the sample in the present study, it is conceivable that the health of the patients included in our study had already suffered considerable
damage by the start of inclusion of this study. It is questionable whether receiving treatment in a care group could still have an impact on that damage, and thus reduce complications and the associated hospitalizations and healthcare costs. This may be easier to achieve for patients who are in a less severe stage of the disease. However, disease management programs have been found to be the most beneficial for patients with poorly controlled diabetes and who have been ill for longer, which means care groups could have a considerable impact on preventing complications in this population.

Secondly, no conclusions can be drawn regarding the effect of care groups versus no care groups on patients’ total healthcare costs, complications and number of hospital days. With the most recent data, it is impossible to make this comparison because almost all patients in the primary care are treated by a care group. Patients not treated by a care group are generally different from patients treated by a care group (for example in a more severe stage of the disease).

Thirdly, in line with the existing literature, some significant factors were identified that influenced patients’ healthcare costs and complications, and SES was included as a variable that was excluded from previous studies. Based on the generally low explained variance of our models, much of the variation is caused by chance or could be explained by unobserved characteristics outside the scope of this study. Many variables can influence the performance of care groups and patient outcomes. Examples of factors that could not be included in this study but could potentially have influenced our outcomes to include ethnicity, duration of the disease, self-reported health status, and level of education. Data on supply variables (such as regional capacity or physicians beliefs) were also unavailable. Some of these variables were partly captured using the variables included in the present study.

Fourth, the aim of the present study was not to explain potential variation between care groups. The aim of this study was to assess whether the variation is observed between care groups. Therefore, several potentially important factors contributing to variation in the care delivery between care groups were not included. For example, data about the maturity of the care group, the number and type of professionals in the care group and details about the care delivery of health care providers included in the care group (in comparison to care from providers not included in a care group). The care groups included in the present study existed for at least six years, but we do not know whether all care groups reached the same level of maturity regarding the provision of integrated diabetes care. It could take over ten years to reach a reasonable mature level of integration and thus the long-term effect of integrated care initiatives may not be visible yet.

Finally, there may have been some underestimation of the number of complications. Since complications were identified based on DTCs, it is possible that some complications were registered under DTCs that were not identified as diabetes-related in this study. In addition, every complication was counted only once, even in cases where a patient developed that complication multiple times. Patients can recover from some complications, and thus they can occur repeatedly. The present study did not capture these, which could also have resulted in underestimation.

Future research may focus on other outcome measures such as patient satisfaction and patient-reported outcome measures (PROMs) in order to broaden the scope of the effect measures. Furthermore, future studies should include more data on participant characteristics and supply variables, data on patients at other stages of the disease and data with a longer follow-up period to detect more complications. Future studies could learn from the analytical approach presented in this study for assessing long-term variation between care groups. If the significant variation is observed, it is important to understand this variation.

The long-term healthcare costs, diabetes-related complications and hospital days of patients with DM2 vary significantly, but care groups seem to account only minimally for this variation. GLMM in combination with year cohorts is a tool to study variations in the long-term outcomes of integrated care initiatives.

Availability of data and material
The individual level data used in this study is not publicly available due to privacy laws.

Code availability
The code used for the analyses is available from the corresponding author on reasonable request.

Declaration of conflicting interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Malou Dorine Menting https://orcid.org/0000-0001-6532-4100

Supplemental material
Supplemental material for this article is available online.

References
1. Zheng Y, Ley SH and Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nature Reviews Endocrinology 2018; 14: 88–98.
2. Struijs JN, Mohnen SM, Molema CCM, et al. Effects of bundled payment on curative health care costs in The Netherlands, RIVM report no. 260013001, 2012, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands.

3. de Bakker DH, Struijs JN, Baan CB, et al. Early results from adoption of bundled payment for diabetes care in the Netherlands show improvement in care coordination. Health Aff 2012; 31: 426–433.

4. Mohnen S, Baan C and Struijs J. Bundled payments for diabetes care and healthcare costs growth: a 2-year follow-up study. The American Journal of Accountable Care 2009; 4: 63–70.

5. Karimi M, Tsiachristas A, Looman W, et al. Bundled payments for chronic diseases increased health care expenditure in the Netherlands, especially for multimorbid patients. Health Policy 2021; 125: 751–759.

6. Struijs JN and Baan CA. Integrating care through bundled payments—lessons from the Netherlands. N Engl J Med 2011; 364: 990–991.

7. Checkland K, Mcldermott I, Coleman A, et al. Complexity in the new NHS: longitudinal case studies of CCGs in England. BMJ Open 2016; 6: e010199.

8. Lewis VA, Colla CH, Carluzzo KL, et al. Accountable Care Organizations in the United States: market and demographic factors associated with formation. Health Serv Rev 2013; 48: 1840–1858.

9. Rocks S, Berntson D, Gil-Salmerón A, et al. Cost and effects of integrated care: a systematic literature review and meta-analysis. Eur J Health Econ 2020; 21: 1211–1221.

10. Baxter S, Johnson M, Chambers D, et al. The effects of integrated care: a systematic review of UK and international evidence. BMC Health Serv Res 2018; 18: 1–13.

11. Struijs JN, de Vries EF, Baan CA, et al. Bundled-payment models around the world: how they work and what their impact has been. The Commonwealth Fund: Issue Briefs 2020.

12. Tsiachristas A, Dikkers C, Boland MRS, et al. Impact of financial agreements in European chronic care on health care expenditure growth. Health Policy 2016; 120: 420–430.

13. Struijs JN, de Jong-van Til JLM, Lemmens L, et al. Bundled payments of diabetes care: effects on care delivery process and quality of care at three-year follow-up, http://www.narcis.nl/publication/RecordID/oai%3Arivm.openrepository.com%3A10029%2F257271 (2012, accessed October 2, 2021).

14. Zonneveld N, Vat LE, Vlek H, et al. The development of integrated diabetes care in the Netherlands: a multiplayer self-assessment analysis. BMC Health Serv Res 2017; 17: 1–9.

15. Dutch Diabetes Federation. Zorgstandaard Diabetes.nl [Treatment Guidelines Diabetes Mellitus], https://zorgstandaarddiabetes.nl/type-2/ (2015, accessed October 2, 2021).

16. Bakker Dd, Raams J, Schut E, et al. Integrale bekostiging van zorg: werk in uitvoering. Eindrapport van de Evaluatiecommissie Integrale Bekostiging [Bundled Payment in Healthcare: work in progress. Final report of the Evaluation Committee Bundeld Payment]. Den Haag: ZonMw, 2012.

17. Klomp M, Romeijnders A, de Braal E, et al. Transparante Ketenzorg Rapportage 2019 Zorggroepen: Diabetes Mellitus, VRM, COPD en Astma. Spiegel voor het verbeteren van chronische zorg [Report on transparency in integrated diabetes care and care groups], 2019; 13–20.

18. Elissen AM, Adams JL, Spreeuwenberg M, et al. Advancing current approaches to disease management evaluation: capitalizing on heterogeneity to understand what works and for whom. BMC Med Res Methodol 2013; 40: 13–40.

19. Campmans-Kuijpers MJE, Baan CA, Lemmens LC, et al. Diabetes quality management in Dutch care groups and outpatient clinics: a cross-sectional study. BMC Res Notes 2014; 7: 1–7.

20. Mohnen SM, Molema CC, Steenbeek W, et al. Cost variation in diabetes care across Dutch care groups? Health Serv Res 2017; 52: 93–112.

21. Eijkenaar F, van Kleef RC, Van Veen SHCM, et al. Onderzoek risicovereveningsmodel 2014: berekening normbedragen [research on risk equalisation system 2014: calculation of marginal costs]. WOR 658, iBMG, Erasmus Universiteit Rotterdam 2013.

22. Freeman J. Management of hypoglycemia in older adults with type 2 diabetes. Postgrad Med 2019; 131: 241–250.

23. Singh SS, Roeters-van Lennep JE, Lemmers RFH, et al. Sex difference in the incidence of microvascular complications in patients with type 2 diabetes mellitus: a prospective cohort study. Acta Diabetol 2020; 57: 725–732.

24. Stedman M, Lunt M, Davies M, et al. Cost of hospital treatment of type 1 diabetes (T1DM) and type 2 diabetes (T2DM) compared to the non-diabetes population: a detailed economic evaluation. BMJ Open 2020; 10: e033231.

25. Grintsova O, Maier W and Mielck A. Inequalities in health care among patients with type 2 diabetes by individual socio-economic status (SES) and regional deprivation: a systematic literature review. Int J Equity Health 2014; 13: 1–14.

26. Lammers M, Vriend S, Tempelman C, et al. Regionale effecten in de risicoverevening [regional effects in the risk equalization system]. WOR 910. Amsterdam SEO Economisch Onderzoek 2018.

27. De Vries EF, Heijink R, Struijs JN, et al. Unraveling the drivers of regional variation in healthcare spending by analyzing prevalent chronic diseases. BMC Health Serv Res 2018; 18: 1–9.

28. Cao P, Wang K, Zhang H, et al. Factors influencing the hospitalization costs of patients with type 2 diabetes. Asia Pac J Public Health 2015; 27: 55S–60S.

29. Bolker BM, Brooks ME, Clark CJ, et al. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 2009; 24: 127–135.

30. Austin PC, Stryhn H, Leckie G, et al. Measures of clustering and heterogeneity in multilevel Poisson regression analyses of rates/count data. Stat Med 2018; 37: 572–589.

31. Bürkner P-C. Brms: an R package for Bayesian multilevel models using Stan. J Stat Softw 2017; 80: 1–28.

32. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2021.

33. Dowle M and Srinivasan A. data.table: Extension of ’data.frame’. R package version 1.14.0. 2021.

34. Nałborczyk L, Batailler C, Loevenburck H, et al. An introduction to Bayesian multilevel models using brms. J Speech Lang Hear Res 2019; 62: 1225–1242.
35. van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. Nature Reviews Methods Primers; 2021. doi: 10.1038/s43586-020-00001-2.
36. Bürkner P-C. Bayesian distributional non-linear multilevel modeling with the R package brms. arXiv 2017: 1705.11123.
37. Abel G and Elliott MN. Identifying and quantifying variation between healthcare organisations and geographical regions: using mixed-effects models. BMJ Quality and Safety 2019; 28: 1032–1038.
38. De Jong JD, Westert GP, Lagoe R, et al. Variation in hospital length of stay: do physicians adapt their length of stay decisions to what is usual in the hospital where they work? Health Serv Res 2006; 41: 374–394.
39. Nakagawa S, Johnson PCD and Schielzeth H. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J R Soc Interface 2017; 14: 1–11.
40. Klomp M, Mutsaerts JF, Rempe J, et al. Denkraam integratie zorgprogramma’s voor chronische aandoeningen [framework integration care programs for chronic diseases]. InEen 2020: 1–17. https://ieneen.nl/wp-content/uploads/2022/04/Denkraam-integratie-zorgprogrammas-juli-2020.pdf (accessed October 2 2021).
41. Vehtari A, Gelman A and Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comput 2017; 27: 1413–1432.
42. Kristensen T, Laudicella M, Ejersted C, et al. Cost variation in diabetes care delivered in English hospitals. Diabetic Med 2010; 27: 949–957.
43. Simmons D, Yu D, Bunn C, et al. Hospitalisation among patients with diabetes associated with a diabetes integrated care initiative: a mixed methods case study. Future Hospital Journal 2015; 2: 92–98.
44. Obinwa U, Pérez A, Lingvay I, et al. Multilevel variation in diabetes screening within an integrated health system. Diabetes Care 2020; 43: 1016–1024.
45. Elissen AMJ, Duimel-Peeters IGP, Spreeuwenberg C, et al. Toward tailored disease management for type 2 diabetes. American Journal of Managed Care 2012; 18: 619–630.
46. Kautzky-Willer A, Harreiter J and Pacini G. Sex and gender differences in risk, pathophysiology and complications of type 2 diabetes mellitus. Endocr Rev 2016; 37: 278–316.