DEPARTMENT OF BUSINESS AND ECONOMICS

SDU
UNIVERSITY OF
SOUTHERN DENMARK

How are you doing? The benefits of informing health care providers on the performance of their patients

By

Line Planck Kongstad, Giovanni Mellace and Kim Rose Olsen

NHE Discussion Papers: No. 4/2019

ISSN 2596-707X

Network on Health Economics
May, 2019

SDU
UNIVERSITY OF SOUTHERN DENMARK

How are you doing? The benefits of informing health care providers on the performance of their patients

By

Line Planck Kongstad[1], Giovanni Mellace[2] and Kim Rose Olsen[3]

1. National Board of Social Services, Denmark
2. Department of Business and Economics, University of Southern Denmark
3. DaCHE, Department of Public Health, University of Southern Denmark

NHE Discussion Papers: No. 4/2019. ISSN 2596-707X

Network on Health Economics
Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55, DK-5230 Odense M
E-mail: Lho@sam.sdu.dk
www.sdu.dk/ivoe

How are you doing? The benefits of informing health care providers on the performance of their patients

By LINE PLANCK KONGSTAD[§], GIOVANNI MELLACE[¶] AND KIM ROSE OLSEN[*]

Between 2011 and 2013, Denmark implemented a program that informed health care providers of their patients' performance. We exploit the exogenous variation in the timing of enrolment into this program to estimate its causal effects on diabetic patients' health outcomes. We find that hospitalizations were reduced by more than 10% and avoidable hospitalizations by more than 15%, an effect comparable in magnitude to other more expensive programs, such as pay-for-performance. (*JEL* I11. I18. H51. C21).

[§] National Board of Social Services, Denmark (lplk@socialstyrelsen.dk) [¶] Department of Business and Economics, University of Southern Denmark (giome@sam.sdu.dk) *DaCHE, Department of Public Health, University of Southern Denmark (krolsen@sdu.dk). Acknowledgments: The authors acknowledge the Association of Danish Regions for an unrestricted grant. Earlier versions of the paper have been presented at various seminars at the following Universities (Free University of Bozen-Bolzano, Italy; University of St. Gallen, Switzerland; Nova School of Business & Economics, Lisbon Portugal) and at the following conferences: 25th European Workshop on Econometrics and Health Economics in Nyborg; iHEA in Boston 2017, IRVAPP Trento, University of Fribourg; NHESG in Odense 2016, and the Research Workshop on Public Policies, Health and Health Behaviors at the Danish Centre for Social Science Research in Copenhagen 2017. We would like to thank the audience at these seminars and conferences for helpful discussions and comments. A special thanks to Pedro Rosa Dias, and Søren Rud Kristensen for their valuable comments as discussants. Previous versions of this paper have circulated under the title "Can the use of Electronic Health Records in General Practice reduce hospitalizations for diabetes patients? Evidence from a natural experiment". Keywords: Disease management programs, health care quality, treatment effects, Difference-in-Differences, non-pecuniary incentives.

I. Introduction

The market for health care is subject to asymmetric and incomplete information. Physicians' treatment choices is substantially affected by incomplete information on the effectiveness of the care they deliver to their patients (L. Berger et al., 2013). Furthermore, physicians might have a different perception on how their patients are doing if they are supplied with information about how their peers' patients are performing. Hence, supplying physicians with market information may have the potential to reduce market failures in health care.

In this study we estimate the causal effects of a non-pecuniary diabetes disease management program that supply general practice clinics (GPs) with detailed information on the clinical performance of the diabetic patients monitored in the clinic and peer comparisons on the average performance of their patients compared to peer GPs at various geographical levels (municipality, regional, and nationwide). We exploit the fact that not all the GPs enrolled at the same time and use a recent matching estimator where we also match on pre-treatment outcomes. This allows controlling for all the unobservable confounders that were already present in the pre-treatment periods. As we will argue in detail below (see Section IV.A), our identification strategy exploits exogenous variation created by incompatibility between some of the IT systems used in Danish GP clinics and the DMP. Using a Difference-in-Differences (DID) approach we obtain very similar results. Our estimated causal effects show that the supply of information to GPs reduced diabetic patients' risk of hospitalizations by more than 10% after one and two years from enrollment. The effect vanishes in 2015 following the program's discontinuation. We find larger effects (more than15% reduction) on hospitalizations that are expected to be avoidable if primary care works well. All results are robust to various definitions of the diabetes population and treatment. Interestingly, the magnitudes of the treatment effects we estimate are comparable to the sizes of those found for disease management programs that include financial incentives and public

reporting.

In this way, our results contribute to studies of more elaborate information-based disease management programs involving public reporting and/or pay for performance. A caveat in this literature is that the programs involve many elements of which the effect cannot easily be disentangled. Examples are studies on cardiac report cards in the US (A. J. Epstein, 2006, R. Steinbrook, 2006) and the UK Quality and Outcomes Framework (QOF) (F. Eijkenaar et al., 2013, M. J. Harrison et al., 2014). In the case of cardiac report cards, the information is made public to patients with an implicit aim of inducing demand-side effects, but it has been questioned whether the positive effects from the program could have been reached using private information only. Kolstad (2013) shows that public reporting does not imply demand-side effects, but that supply-side effects arise because of physicians' intrinsic motivation to supply good quality of care. Actually, Kolstad argues that his results indicate that. " … *to the extent that information about peers alters surgeons' intrinsic incentives, public release is of less relevance. In fact, contrary to current efforts to simplify provider report cards, it may be preferable to deliver data with more clinical detail* ". Our results support this argument, as they show that effects can occur in a case where information is only private.

Our results also contribute to the literature on the effect of disease management, with pecuniary incentives (M. Dusheiko et al., 2011, W. Han et al., 2016, E. Iezzi et al., 2014, M. Reed et al., 2013). Interestingly, the effect on hospitalization has been found to be present in programs with both high- and low-powered incentives (e.g. financial vs. information sharing alone). For example, the UK QOF, which was implemented in the GP sector involving public information sharing and strong financial incentives to physicians, has been shown by Harrison. et al. (2014) to reduce ACSC hospitalizations by around 10%. In comparison (M. Reed, J. Huang, R. Brand, I. Graetz, R. Neugebauer, B. Fireman, M. Jaffe, D. W. Ballard and J. Hsu, 2013) studies a non-pecuniary scheme

without public information that was implemented in California and find that EHR participation reduce ACSC hospitalizations by approximately 10%. – which is comparable to the findings in (M. J. Harrison, M. Dusheiko, M. Sutton, H. Gravelle, T. Doran and M. Roland, 2014) on the effectiveness of the QOF. This may indicate that non-financial incentives may have an effect comparable to financial incentives in the case of general practice. Our results support this by showing comparable magnitude of effects on hospitalizations in a non-pecuniary program based solely on private information.

The paper is structured as follows. First, section II discusses the conceptual framework for the hypothesis of an effect on physician behavior of private information on how their patients are doing. Then, we describe the main structures of the Danish health system followed by a detailed description of the disease management program in section III. Section IV describes the dataset used, and section V discusses and compares the identification strategies. Results, robustness checks, and discussions are reported in section VI. Finally, section VII concludes.

II. Conceptual framework

Physicians' behaviors, when operating as self-employed entities - as is the case for GPs in Denmark - have traditionally been modeled as based upon maximization of a utility function where the decision of supply of services q_i, is made as a tradeoff between income and leisure (M.; Pauly Gaynor, MV, 1990, T. Iversen, 2004, K. R. Olsen, 2012, A. Scott, 2000). The DMP studied in this paper involves peer comparison and information on patient treatment quality, which the physician may not have full information about. The theoretical background for the hypothesis that these two elements may change physician behavior is an assumption of non-pecuniary utility entering the physician's utility function in addition to income and leisure. The non-pecuniary utility component is related to physicians benefits from improving patients health – also referred to as provider

altruism (M. Chalkley and J. M. Malcomson, 1998, R. P. Ellis and T. G. McGuire, 1986, J.T. Kolstad, 2013, A; Mak Ma, H., Forthcoming).

(J.T. Kolstad, 2013) defines a model that nicely shows how the two components of the Danish DMP (information and peer comparison) affect physicians' utility and treatment decisions. Kolstad's' model predicts that a physician is willing to invest in quality and effort beyond an equilibrium defined by marginal revenue equalizing marginal costs because she gains utility from increasing the health of the patient. As there is uncertainty about the health of the patient and how the level of quality supplied differ from peers he shows that increased market information changes the shape of the utility components and hence the choices by the physician.

Our setting differs from the cardiac surgeon case that Kolstad considered. First of all GPs are self-employed which involves that leisure should be added in the utility function. Hence, the level of effort chosen by the physician is not only traded off against profit and intrinsic utility but also leisure. Second, in the DMP we consider, the market information is not shared with the patients. Therefore, no change in demand from prom patients as a consequence of increased information can be expected. Third, the level of competition among GPs is small because patients are listed with a certain GP, and there are regulated restrictions as well as costs related to changing from one GP to another. Therefore we do not expect that demand for services are affected by other GPs quality to the same extend, as is the case for cardiac surgery. We can assess if these deviations from the cardiac surgeon case give raise to any change in the predictions of the model by adding leisure and neglecting demand effects. First let us define the utility of GP *i* as defined by J.T., Kolstad, 2013

$$U_i = \Pi_i(\theta_i, \theta_{-i}|\Omega) + \Gamma_i(\theta_i, \theta_{-i}|\Omega) + \Lambda_i(\theta_i)$$

where $\Pi_i(\theta_i, \theta_{-i}|\Omega)$ describes the utility component from profit, which in the case of demand side effect can be defined as $\Pi_i(\theta_i, \theta_{-i}|\Omega) = \mathrm{q}_i(\theta_i, \theta_{-i}|\Omega)p - c(\mathrm{q}_i, \theta_i)$. Where p refers to fixed prices and c to the cost. θ_i is the quality level chosen by the GP and θ_{-i}, θ_j are the quality levels of all peers and the peers that the GP is compared with, respectively. Ω defines the information available in the market. In our case Ω consist of the clinical performances of the GPs own diabetic patients as well as that of her peers. $\Gamma_i(\theta_i, \theta_{-i}|\Omega)$ defines intrinsic utility and $\Lambda_i(\theta_i)$ defines utility from leisure and is assumed decreasing in θ_i. The first order condition (FOC) for the GPs choice of θ_i is derived as

$$(\partial \mathrm{q}_i(\theta_i, \theta_{-i}|\Omega)/\,\partial\theta_i)p + \partial\Gamma_i(\theta_i, \theta_{-i}|\Omega)/\,\partial\theta_i = \partial \mathrm{c}\left(\theta_i, \theta_j^*\right)/\,\partial\theta_i - \partial\Lambda_i(\theta_i)/\,\partial\theta_i$$

The FOC states that the level of quality is chosen such that the marginal revenue and the marginal gain in intrinsic utility equals the sum of the marginal costs and the loss of leisure. The leisure component will all else equal involve a lower quality level but as it has been defined is does not change the effect of market information. Now, neglecting the demand side effect of market information as well as direct competition on quality involve profit to be defined as independent of market information and peers level of quality: $\Pi_i(\theta_i) = \mathrm{q}_i(\theta_i)p - c(\mathrm{q}_i, \theta_i)$ and hence the FOC to be defined as

$$(\partial \mathrm{q}_i(\theta_i)/\,\partial\theta_i)p + \partial\Gamma_i(\theta_i, \theta_{-i}|\Omega)/\,\partial\theta_i = \partial \mathrm{c}(\theta_i)/\,\partial\theta_i - \partial\Lambda_i(\theta_i)/\,\partial\theta_i$$

which show that changes in market information and peer comparison now only works through the intrinsic utility component. The prediction of the model should still be the same although the effect may diminish without demand side effects and with an additional trade off against leisure. However, as the empirical results in Kolstad, (2013) suggest demand side effects to be small – the difference in context may not matter a lot after all.

III. Context - the Danish health system and the disease management program

A. Danish health care

Denmark, like other Scandinavian countries, is characterized by having a tax-financed health care sector with universal coverage and a focus on equal access for all (C.H.; Christiansen Lyttkens, T.; Häkkinen, U.; Karboe, O.; Welander, A. , 2016). Hospitals are almost entirely public, and private hospitals are mainly specialized in selected surgeries. Primary care is characterized by self-employed general practitioners (GPs) acting as gate-keepers to specialized care (K.R. Olsen et al., 2016). The remuneration scheme is mixed, with a 70% fee for service and 30% capitation. For many years, there has been scarcity of GP physicians – meaning that most GPs have the number of patients they prefer and that competition for patients is limited. In 2018 more than half of all GP practices have closed their list – meaning that they are not taking in new patients. Furthermore, patients' mobility is rather low and there are restrictions in the sense that a patient must choose a GP within a certain distance from his or her home.

B. The diabetes disease management program

The term disease management program (DMP) has been used in many different contexts and there exists several definitions. (G. Ellrodt et al., 1997) defines disease management as programs that use a systematic approach to care and included more than 1 intervention component. (K. Knight et al., 2005) reviewed the DMP literature on diabetes care based on the definition in (G. Ellrodt, D. J. Cook, J. Lee, M. Cho, D. Hunt and S. Weingarten, 1997, G. Schrijvers, 2009) and assess the following components: guidelines; protocols; algorithms; care plans; and systematic patient or education programs. As related to disease management (G. Schrijvers, 2009) define chronic disease management (CDM) as consisting of a group of coherent interventions designed to prevent or manage one or more chronic conditions using a systematic multidisciplinary approach potentially

employing multiple treatment modalities.

In this section we will carefully explain the content of the disease management program we study in this paper and refer to it as diabetes disease management program based on electronic health records, private feedback reports, peer comparison, and with absence of direct pecuniary incentives.

The diabetes DMP that we consider in this study was introduced as a pilot study/development project with a limited number of participants in the period 2007-10. Then, in the National contract agreed between the Association of General Practice and the Association of Danish Regions in December 2010, participation was made mandatory and GPs were obliged to enroll in the program within three years – i.e. by the end of 2013 at the latest. Enrolment did not necessarily mean that the activities in the DMP were implemented and some may have enrolled without actually implementing the program.

The program was abandoned in august 2014 due to problems with the legislative approvals of some of the data recorded at patient level as part of the DMP. Hence, by looking at outcomes including 2015 we can establish whether participation up until the first 8 month of 2014 had effects on hospitalizations in 2015.

The DMP has been described in a couple of studies (L. Lippert, 2014, H. Schroll et al., 2012). but the main content will be explained here as well. The program was based on electronic health records (EHR) with electronic registration of symptoms and/or diagnoses related to each patient contact. The International Classification of Primary Health Care (ICPC) was used as coding system[1]. Furthermore, laboratory test and services supplied were gathered in a database and used to construct easy access private feedback reports of treatment for diabetic patients. Unfortunately, these patient level data are not available for research.

[1] http://www.kith.no/upload/2705/icpc-2-english.pdf

The potential benefits of introducing EHR are well known from the literature on the US HITECH Act; often divided into the effect of implementing population registries, and the effect of clinical decision support (V. Patel et al., 2015). The Danish program includes both components.

The EHR system increases the information available at the point of care and allows the GP to plan better monitoring of the patients. For example, the EHR system allows the GP to obtain an overview of diabetic patients that have not had their annual control or who were not performing well on certain clinical indicators. Figure 1 show an example of the information that was supplied as part of the DMP.

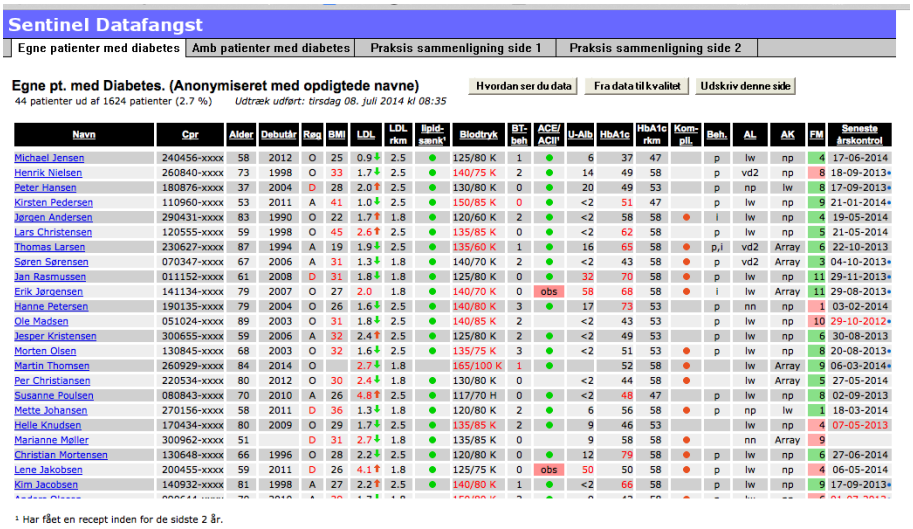


FIGURE 1: POPULATION OF DIABETIC PATIENTS

Notes: Screen dump from the demo version of the EHR system. Names and id-numbers are fictitious. The figure show that the new DMP allowed for easy access to overviews of the diabetic population and their performance on important clinical parameters as e.g.: HbA1c levels; Blood pressure; Usage of lipid lowering drugs; BMI; Smoking status. GPs report that they used the overview of diabetic patients' performances to pro-actively invite patients with critical values for a visit (Lippert. 2014).

The information in figure 1 informs the physician about the current health state of the patients and how they deviate from preferred clinical outcomes. As discussed in the section of conceptual framework, this information may incentivize the GP to change the level of services supplied to the patients.

The program also supplied private peer comparisons to the GPs about their own performance as compared to their peers at the municipality, regional, and national level. Again, our conceptual framework explained how peer comparison may impact GPs behavior. Figure 2 illustrates the content of these peer comparisons.

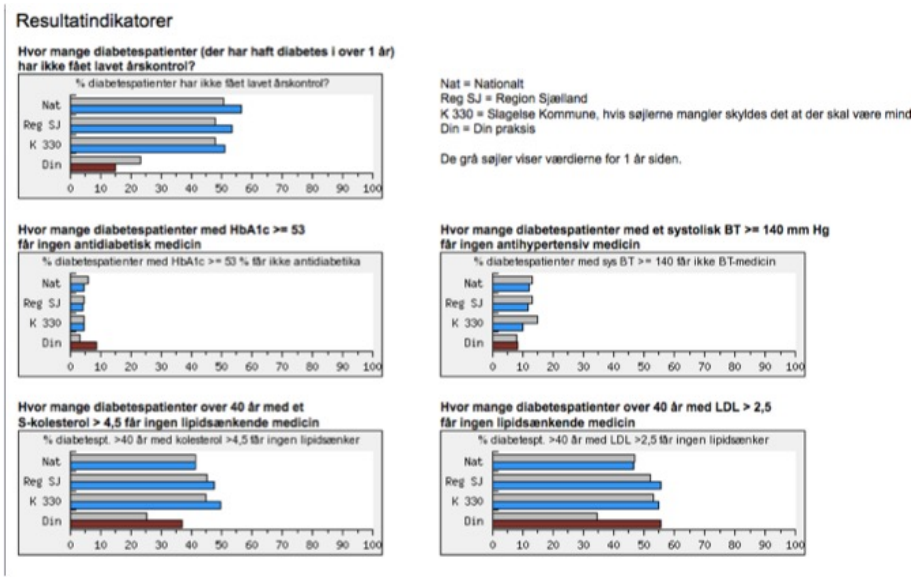


FIGURE 2: RELATIVE PERFORMANCE CHART COMPARED TO PEERS

Notes: With the new DMP the treated GP can see how their diabetic patients performed compare to the average in the country. region. and municipality. The graphs show selected performance measures for the GPs own clinic as compared to all clinics in Denmark. all clinics in the region of location. and all clinics in the municipality of location. The top left graph shows the share of diabetes patients that have had an annual control. The graph in the middle left hand side shows how many patients with HbA1c >= 53 do not receive any diabetes drugs. The lower left-hand graph shows how many diabetes patients above 40 years of age have S-cholesterol levels above 4.5 but does not receive any lipid lowering medication. The top right-hand graph shows the number of patients with risk of hypertension who do not receive anti-hypertension medication. The lower right-hand graph shows how many patients above 40 years of age with an LDL level higher than 2.5 do not receive lipid lowering medications.

IV. Data

A. Definition of treatment and control groups

Although participation was made mandatory in 2011, no enforcement mechanisms existed, and GPs were given 3 years to enroll. Installing the EHR and integrating it with the clinic's IT system was the first step. A review concluded that technical issues represented one of the main obstacles to implementation (C. A. McGinn et al., 2011). Hence, we expect that differences in differences in the

integration ability between the IT system used (11 different systems are used in Denmark) and the EHR technology generated some exogenous variation in the timing of enrolment. In fact, some of the systems could not be immediately integrated with the EHR. Given that there is no reason to believe that using a specific IT system is related to the quality of disease management and therefore hospitalizations, this is likely to create exogenous variation in EHR enrolment. Unfortunately, since we do not observe which IT system is used by the clinic, we cannot further exploit this exogenous variation in an IV setting.

Notice that, enrolment does not necessarily mean implementation. Moreover, mere installation of the EHR system does not in itself improve treatment quality. Therefore, we need to take this into account in our treatment definition. Participation in the program is defined using data from the Danish Quality Unit of General Practice (DAK-e). We observe two indicators of implementation: i) monthly percentages of visits in the clinic that were registered in the EHR system; and ii) the dates in which the GP logged into the diabetes population overview (figure 2). We use this information to define our treatment as an indicator of actual implementation of the DMP rather than just enrolment. Hence, we define a binary treatment variable which is equal to one if in 2012 the median monthly EHR usage reaches at least 70% of patient visits coded, without reaching this threshold in any year prior to 2012 and with a further condition that the GP has logged into the diabetes population overview at least once. Reed et al (2013) use an 80% threshold to define their EHR implementation indicator. As a 70% threshold was explicitly used by the GPs themselves as an indication of implementation (F. Ulstrup, 2012) we use this threshold in our main specification, our results are robust to using 60% and 80% thresholds instead. We will focus on 246 GPs who are treated according to our definition in 2012. The reason is that only a very small number of

extremely selected clinics enrolled before 2012. The first movers[2] in 2011 consist of about 60 GPs who likely include GPs who have a special interest in the DMP and therefore were more selected. We do not look at GPs who enrolled in 2013 or 2014 as we can only look at 1 or 2 years of outcomes after enrolment for those cohorts.

As the number of enrolled GP's increases over time there is a risk of having too few control GP's in the later years. Hence. the control group at time *t* consists of all GPs that were either never enrolled in the program at time *t* or were enrolled but without ever coding more than 10% of their visits and without using the population overview (figure 2). Under the mild assumption that slightly using the DMP was either beneficial or had no impact for those GPs. the magnitudes of our effects are, if anything, conservative. The 10% threshold is chosen to make sure that the number of coded visits in the control clinics was likely not enough to cover their entire diabetic population.

To avoid including GPs with very few diabetic patients, we restrict our analysis to GPs with more than 20 diabetic patients. We further drop GPs that are not present in all years of observation. This definition of control GPs leaves us with 421 control GPs until 2012. 325 in 2013, and 317 in 2014 and 2015.

B. Definition of the diabetes population

We identify diabetic patients using an algorithm suggested by the national Danish diabetes register, which is comparable to the algorithm suggested by Iezzi et al 2014. Patients are considered to be diabetic if they had either more than two HbA1c or blood glucose measurements within a year provided by the GP or a specialist, or a prescription of diabetes medications, or were hospitalized because of diabetes. As diabetes is a chronic disease, we assume that a patient defined as diabetic in

[2] For this population we find a significant decrease in total hospitalization three and four years after enrolment and no effects on the other two hospitalization measures. More detailed results are available from the authors upon request.

one period will remain diabetic in all subsequent periods.

We make two restrictions on the diabetes population to avoid composition bias and bias from the risk that patients actively select GPs that enroll in the program. First, there is a risk of composition bias if enrolment in the DMP involves a systematic increase in the number of a certain type of patients defined by the diabetes algorithm. If this is the case, changes in average hospitalization rates at GP level may be due to the change in the composition of the diabetic patients' population rather than an effect of the DMP. An obvious example would be when less severe diabetic patients are more likely to be detected by treated GPs because of their extra awareness. This would mechanically reduce the average hospitalization rate for the treated clinics and this reduction would not be due to an increase in treatment quality. To avoid this issue, we only consider diabetic patients who are defined as diabetic patients before 2011. As a robustness check, we also run our analysis on the unrestricted population of diabetic patients. To avoid patients' selection into treated GPs, we further restrict the population to diabetic patients who are listed with the same GP from 2011 to 2015.

C. Control variables

Although the program became mandatory in 2011, the time of enrolment was largely voluntary until the end of 2013. Furthermore, enrolment was not necessarily followed by implementation of the program. Thus, our identification strategy must deal with potential self-selection of GPs into the program. Not only are we controlling for a rich set of control variables, but we also use methods that control for unobserved confounders.

We argue that the selection bias would mostly be induced by differences between treated and control GPs in their interest in diabetes, practice size, treatment pattern, IT knowledge and the socio-demographics of their diabetic patients. We provide below a detailed explanation for each of

these potential selection bias channels

Interest in diabetes treatment: Because the program was developed around diabetes treatment, we believe that the first movers may have a specific interest in diabetes. GPs with special interest in diabetes are also expected to do better in terms of treatment quality and hence hospitalizations, even without participation in the program. Because this difference in performance is likely to be present in the pre-treatment period, our main identification strategy will capture this by including pre-treatment outcomes in our set of control variables. When using Difference-in-Differences we will assume that the difference in interest in diabetes among treated and non-treated GPs does not change between the pre- and the post-treatment periods.

Practice size: Several studies claim that practice size affects EHR adoption and health outcome (D. Gans et al., 2005, W. Han, R. Sharman, A. Heider, N. Maloney, M. Yang and R. Singh, 2016, J. D. Ketcham et al., 2007, Y. Wang et al., 2006). Hence, we expect participating GPs' practices to be larger than non-participants' practices. Practice size can also have impact on hospitalization. Fortunately, practice size as by the number of patients is observed in our data and included as a control variable.

Treatment pattern: Information on treatment pattern is based on data from the Danish National Health Service Register (NHSR), which contains information about the activities of health professionals contracted with the tax-funded public healthcare system (including GPs). Because GP's are partly paid by fee for services, we can use the fee structure to give an indication of the treatment pattern of the GP. We include the following treatment pattern variables for the GPs' diabetic patients: number of visits, number of telephone visits, number of e-mail visits, number of blood tests, number of planned annual control visits, use of diabetes medication, influenza vaccinations.

IT knowledge: Because the program is based on an IT solution, participants may be more skilled or better organized to deal with new technology. The differential in the attitude towards new technology is likely to be present in the pre-treatment periods already and captured by including the pre-treatment outcomes. Moreover, we include the number email and telephone consultations to proxy for IT skills and propensity to use new technologies.

Characteristics of the patients on the list: Because the implementation of a DMP requires a certain amount of excess capacity, it may be that GPs with less deprived or frail patients may be more likely to participate. We control for several patient's characteristics. First, we use the Charlson Comorbidity index as a measure of the level of comorbidities for diabetic patients. We further use variables included in the Danish Deprivation Index (DADI) (see e.g. (K. R. Olsen, 2012)). These variables are based on rich patient level socio-demographic characteristics, provided by Statistics Denmark, and measure the share of diabetic patients that belong to the following categories: 1) unemployed between 20 and 59 years of age, 2) between 25 and 59 years of age without secondary education, 3) between 25 and 65 years of age with low income, 4) between 18 and 59 years of age on public benefits, 5) with non-Western ethnicity, 6) above 30 years of age who live alone and 7) above 70 years of age with a low level of disposable income. We further control for regional dummies.

Table 1 shows the difference in mean observable characteristics between treated and control GPs in 2010. As expected, treated and control GPs differs with respect to several of those variables. The treatment groups have significantly larger practices, and more consultations and prophylactic visits for their diabetic patients. We do find some differences in patient socio-demographic characteristics. Patients of treated GPs seem to have slightly lower morbidity, measured by the Charlson index, than patients of GPs in the control group. We do not find any significant regional variations in enrolment.

TABLE 1: CO-VARIATES BY TREATMENT AND CONTROL IN 2010

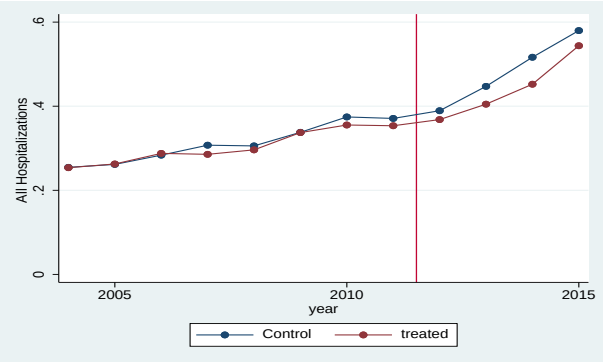
	Treated N=246	Control N=421	Difference	
Planned preventive visits	1.43	1.26	0.16	**
GP Visits	6.90	6.73	0.17	
Influenza vaccinations	0.30	0.31	-0.01	
Blood tests	1.27	1.35	-0.08	
Email consultations	0.81	0.72	0.09	
List size	2353.20	1932.90	420.30	***
Charlson index	0.47	0.53	-0.05	***
Diabetes drugs	992.87	1051.80	-58.93	*
Telephone consultations	4.84	5.45	-0.61	***
DADI_1	0.04	0.04	0.00	
DADI_2	0.29	0.28	0.01	
DADI_3	0.09	0.10	-0.01	**
DADI_4	0.12	0.10	0.02	***
DADI_5	0.09	0.11	-0.02	**
DADI_6	0.30	0.32	-0.02	***
DADI_7	0.25	0.27	-0.02	***
Central Denmark Region	0.21	0.25	-0.04	
Region of Southern Denmark	0.26	0.21	0.04	
Capital Region of Denmark	0.32	0.30	0.02	
Region Zealand	0.12	0.16	-0.04	

Notes: All variables are measured in 2010 as averages per diabetes patient per GP clinic except the Regional dummies which are measured at the GP level. We reports averages for 2012 control group. similar results are found for the remaining control groups. The reference region is North Denmark Region. DADI_1: unemployed between 20 and 59 years of age. DADI_2: between 25 and 59 years of age without secondary education. DADI_3 between 25 and 65 years of age with low income. DADI_4: between 18 and 59 years of age on public benefits. DADI_5 with non-Western ethnicity. DADI_6 above 30 years of age who live alone. and DADI_7 above 70 years of age with a low level of disposable income.

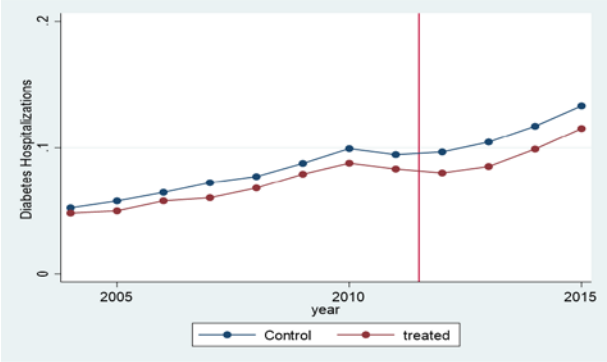
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

D. Hospitalization data

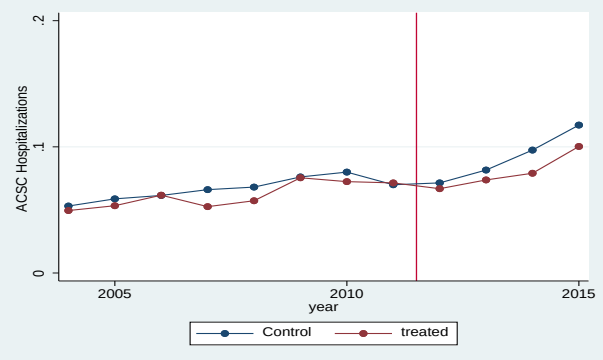
Diabetic patients are known to have severe comorbidities. For this reason, arguably, total hospitalization is the most informative among our outcomes of interest. A UK study, for example, reports that more than 85% of diabetic patients suffer from at least one additional chronic disease (P. Kasteridis et al., 2015). The population overview (figure 2) also supplies information on clinical performance on parameters related to other morbidities. e.g. cardiovascular diseases. Hence, the program should reduce both diabetes related hospitalizations as well as hospitalizations related to other factors.



3.A: ALL HOSPITALIZATIONS.



3.B: DIABETES HOSPITALIZATIONS.



3.C ACSC HOSPITALIZATIONS.

FIGURE 3: COMMON TREND IN OUTCOMES

Notes: The hospitalization variables are measured as averages per diabetes patient per GP clinic. Average of all hospitalizations, diabetes hospitalizations, and ACSC hospitalizations in 2010 are 0.41, 0.09 and 0.11 respectively.

Most of the related studies use ACSC hospitalizations as outcome. M. Reed et al (2013) includes ACSC hospitalizations related to diabetes and cardiovascular disease. E. Iezzi et al (2014) only include diabetes ACSC's and M. Dusheiko et al (2011) only look at ACSC diabetes admissions coded as emergency admissions. Based on the above argument of severe comorbidities we use total ACSC admissions. Like the above-mentioned studies, we use the AHQR definition of ACSC hospitalizations (Agency for Healthcare Quality and Research (AHQR), 2001). The information on hospitalizations is obtained from the Danish National Patient Register covering all somatic hospital treatments.

Figure 4 shows the trends in the mean of our hospitalization outcomes for treated and control GPs. Our main identification strategy does not rely on a common trend assumption, but we do not find any evidence of violations of the common trend assumption in the pre-treatment period for all our outcomes (see also the falsification test in the results tables below).

V. Identification strategy

A. Identification with panel data

Let us start by introducing some notation. We denote by D_i our treatment indicator, which is equal to 1 if GP i participates in the DMP. We denote by $Y_{i.t}$ the observed outcome of GP i at time t and by $X_{i.t}$ our set of GPs and patients' observable characteristics. We denote by $Y_{i.t}^1$ and $Y_{i.t}^0$ the potential outcomes that GP i would achieve at time t with and without participation in the program. Under the stable unit treatment value assumption (SUTVA), one of the two potential outcomes is observed for each GP, according to their treatment status, as described in the following observational rule:

$$Y_{it} = D_i Y_{i.t}^1 + (1 - D_i) Y_{i.t}^0$$

We are interest in the Average Treatment Effect on the Treated (ATET), which is the only parameter DiD can identify without imposing very restrictive assumptions. In our setting the ATET measures the effect of the program on GPs who participated in it, and it is arguably of great policy interest.

The ATET at time t is defined as:

$$ATET_t \equiv E(Y_{i.t}^1 - Y_{i.t}^0 | D_i = 1)$$
$$= \underbrace{E(Y_{i.t} | D_i = 1)}_{Identified} - E(Y_{i.t}^0 | D_i = 1)$$

The first term of the $ATET_t$ is identified by the observable mean outcome of the treated GPs. The second term requires the counterfactual mean potential outcome treated GPs would have experience had they not participated in the program, which is unobservable the treatment period. To identify the $ATET_t$ in the post-treatment periods we need to impose assumptions that enable us to express $E(Y_{i.t}^0 | D_i = 1)$ in terms of observable quantities.

B. Matching with pre-treatment outcomes

We first consider a conditional independence assumption (CIA), where we also include in the conditioning variables pre-treatment values of the outcomes.

Assumption CIA (Conditional Independence)

$$E(Y^0_{i.k+\tau}|D_i = 1.X_{i.k} = x.Y_{i.k-1} = y_{k-1}\,.Y_{i.k-2} = y_{k-2}\,.\ldots.Y_{k-T} = y_{k-T})$$
$$= E(Y^0_{i.k+\tau}|D_i = 0.X_{i.k-1} = x.Y_{i.k-1} = y_{k-1}\,.Y_{i.k-2} = y_{k-2}\,.\ldots.Y_{k-T} = y_{k-T})$$

Where k is the period in which the treatment starts and $Y_{k-1}.Y_{k-2}\ldots..Y_{k-T}$ are the observed pre-treatment outcomes. CIA assumes that conditional on the pre-treatment outcomes and the other control variable X, there is no selection bias. It is important to stress the fact that including pre-treatment outcomes allows us to control for all the unobservable confounders that were also present in the pre-treatment periods. Notice that as we control for pretreatment values of the outcome we are implicitly assuming that there are no anticipation effects. To make sure that for each treated unit there exist a comparable control unit we also need to impose the following common support assumption:

Assumption CSM (Common Support Matching*)*

$$\mathrm{P_x} \equiv \Pr(D_i|X_{i.k} = x.Y_{i.k-1} = y_{k-1}\,.Y_{i.k-2} = y_{k-2}\,.\ldots.Y_{k-T} = y_{k-T}) < 1.$$

$$\forall\, x \in supp(X)\ and\ \forall\, y_{k-1.}\ldots.y_{k-T} \in supp(Y)$$

Notice that CSM allows for the propensity score $\mathrm{P_x}$ to be zero as we only focus on the ATET.

It is easy to show that under CIA and CSM. we can identify the ATET at any period $t \geq k$ as:

$$ATET_t = E(Y_{i.t}|D_i = 1) - E(E(Y_{i.t}|X_{i.k} = x.Y_{i.k-1} = y_{k-1}\,.Y_{i.k-2} = y_{k-2}\,.\ldots)|D_i = 1)$$

The fact that this expression only depends upon observable variables implies that the $ATET_t$ is identified. In practice, the $ATET_t$ can be estimated by any matching estimator. The main idea of this class of estimators is to compare each treated unit with control units that are similar in terms of the covariates (including the pre-treatment outcomes). We use the distance weighted radius matching estimator of (Martin Huber et al., 2015)), which is suggested in the simulation study of (Martin Huber et al., 2013). Following (P. Rosenbaum and D. Rubin, 1983), who show that one can use the propensity score as the only conditioning variable, this estimator compares treated and control with similar values of the propensity score P_x. Furthermore, to make sure that there are no imbalances in strong predictor of the outcome that would lead to a big bias (see (Martin Huber, Michael Lechner and Andreas Steinmayr, 2015)) we use list size, diabetes drugs, the regional dummies, the most recent value of the outcome for the Mahalanobis distance. Finally, following (A. Abadie and G.W. Imbens, 2006, D. Rubin, 1974) we use the standard regression-based bias correction.

C. Comparison of DiD and Matching

Instead of matching on pre-treatment outcomes (simply matching), one could also use Difference-in-Differences (DiD). Let $t = k$ be the last pre-treatment period (2011 in our data); together with Assumption CSM (excluding the pre-treatment outcomes). DiD imposes the following common trend assumption:

Assumption CT (Common Trend)

$$E(Y^0_{i.k+\tau}|D_i = 1.X_{i.k} = x) - E(Y^0_{i.k}|D_i = 1.X_{i.k} = x) =$$

$$E(Y^0_{i.k+\tau}|D_i = 0.X_{i.k} = x) - E(Y^0_{i.k}|D_i = 0.X_{i.k} = x). \forall\, x \in supp(X)$$

Assumption CT is equivalent to assuming that, conditional on the observable covariates, the

selection bias is constant over time. It is easy to see that the $ATET_{k+\tau}$ is then identified as:

$$ATET_{k+\tau} = E(Y_{i.k+\tau}|D_i = 1) - E(E(Y_{i.k+\tau}|D_i = 0.X_{i.k} = x)|D_i = 1)$$
$$- E(Y_{i.k}|D_i = 1) - E(E(Y_{i.k}|D_i = 0.X_{i.k} = x)|D_i = 1)$$

Using panel data, this value can be estimated using the same matching estimator (without the pre-treatment outcomes) using the first difference $\Delta Y_{k+\tau} = Y_{k+\tau} - Y_k$ instead of $Y_{k+\tau}$ as an outcome.

Both (G. Imbens and J. Wooldridge, 2009) and (M. Lechner, 2015) argue that matching on pre-treatment outcomes is preferable to DiD. The two identification strategies, however, are based on assumptions which are not nested, in the sense that in general, the violation of one does not exclude the other.

DiD allows for the presence of selection bias as long as it is time constant and there are no anticipation effects; whereas matching allows for anticipation effects but assumes zero selection bias conditional on the observed covariates and on pre-treatment outcomes. Only when the selection bias is zero in both pre- and post-treatment periods do the two approaches lead to the same results.

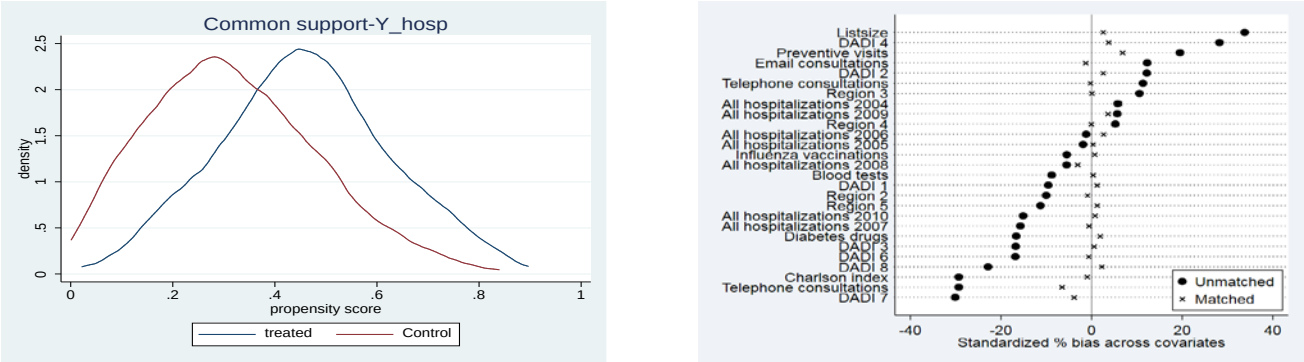
DiD allows for time-invariant unobserved confounders. Matching, by contrast, requires all of the unobservable confounders to be already present in the pre-treatment periods. (S. Chabé-Ferret, 2015) provides a simple model in which matching is consistent if the selection bias is due to transitory shocks only, whereas DiD is consistent if the selection bias is due only to permanent individual fixed effects. He also shows that it is not possible to combine the two approaches to get rid of both sources of bias. In fact, if one tries to condition also on pre-treatment outcomes in a DiD, the resulting estimator is only consistent under the same CIA assumption imposed by matching. (S. Chabé-Ferret, 2015) also shows that conditioning on several pre-treatment outcomes might help to

reduce the bias of matching in the presence of permanent individual fixed effects. Given the richness of our data and the consideration we made in section II, matching is preferable in our setting. In fact, we arguably observe many GPs' and patients' characteristics as well as 7 years of pre-treatment outcomes, which make the CIA likely to hold.

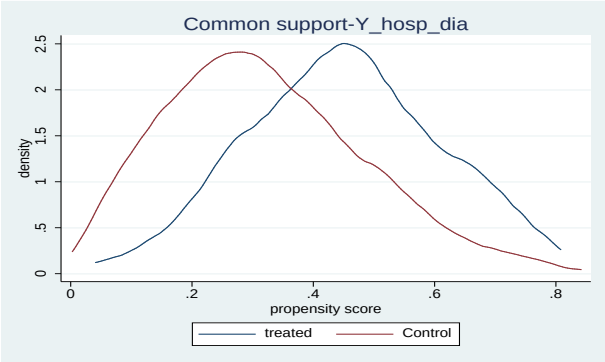
VI. Results

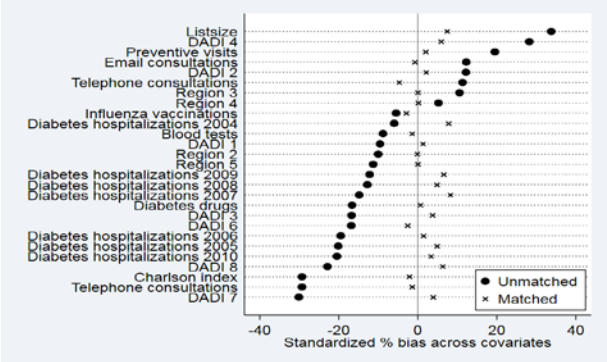
A. Matching quality

We first estimate the propensity scores (one for each outcome) using a Probit model. The three upper panels of figure 4 show the overlap in the propensity scores for total hospitalizations, diabetes hospitalizations and ACSC hospitalizations respectively. The overlap is good and there appear to be no common support issues. Looking closely at potential support issues the maximum number of observations off-support is 2 and in most cases is only one. We therefore omit the support tables. The bottom panels of Figure 4 report the reduction in bias after matching for each of the outcome variables.

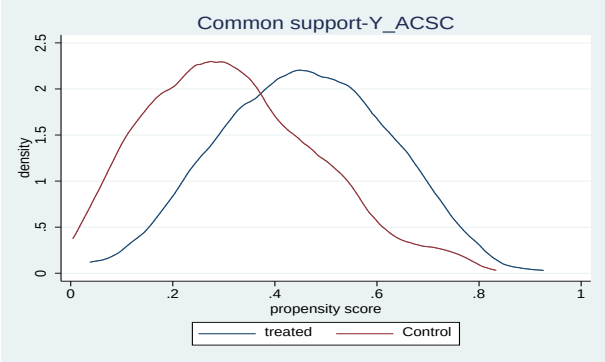


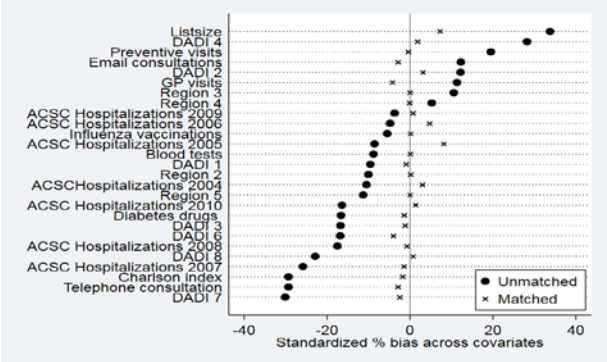
4.A: ALL HOSPITALIZATIONS





4.B: DIABETES HOSPITALIZATIONS.





4.C: ACSC HOSPITALIZATIONS

FIGURE 4: COMMON SUPPORT AND BIAS REDUCTION

Notes: All variables are measured in 2010 as averages per diabetes patient per GP clinic except the Regional dummies which are measured at the GP level. DADI_1: unemployed between 20 and 59 years of age. DADI_2: between 25 and 59 years of age without secondary education. DADI_3 between 25 and 65 years of age with low income, DADI_4: between 18 and 59 years of age on public benefits. DADI_5 with non-Western ethnicity, DADI_6 above 30 years of age who live alone, and DADI_7 above 70 years of age with a low level of disposable income. We used the psmatch2 Stata library of Edwin Leuven and Barbara Sianesi to create the bias reduction graphs.

The graphs above show a good overall reduction in the standardized bias. We do not find any statistically significant difference between our control variables after matching (see Tables A4 to A6 in the appendix).

B. Main results

Table 2 shows the results of our main specification. We find a strong and statistically significant reduction in total hospitalizations. In 2010 the average annual hospitalization rate in treated clinics was 0.40. Therefore, our results indicate a reduction of 9% one year after enrolment and of 11% after two years. The effect seems to vanish once the program is abolished in 2015, however this

result is only an indication, as it does not consider possible endogeneity of the abolishment of the program. We also find strong reduction in both ACSC and diabetes related hospitalizations. Annual ACSC hospitalization rates, which had an average of 0.08 in 2010 for the treated clinics, are reduced by 13.75% and 16.25% two and three years after enrolment, respectively. One year and two years after enrolment, annual diabetes hospitalization rates are reduced by 12% with respect to their average of 0.10 in 2010 for the treated GPs. The estimated effect is similar in the second year after enrolment but not statistically significant and totally disappear in 2015 – after the program had been abolished.

We ran a standard falsification tests using the four closest pre-treatment years as an outcome. Because the true treatment effect is necessarily zero (assuming no anticipation effects) in the pre-treatment periods, PSM estimates only the selection bias in those periods, which must be zero under our assumptions. All the estimated selection biases are insignificant as expected.

TABLE 2 ATET ESTIMATE – MAIN SPECIFICATION

	Falsification				ATET			
	2008	2009	2010	2011	2012	2013	2014	2015
Hospitalizations	-0.005	0.002	-0.007	-0.004	-0.014	-0.037**	-0.043*	-0.015
P-value	*0.522*	*0.844*	*0.547*	*0.720*	*0.352*	*0.040*	*0.055*	*0.461*
ACSC hospitalizations	-0.006	0.003	-0.005	0.004	-0.002	-0.011*	-0.013*	-0.006
P-value	*0.157*	*0.496*	*0.365*	*0.418*	*0.687*	*0.074*	*0.059*	*0.405*
Diabetes hospitalizations	-0.001	-0.002	-0.002	0.001	-0.002	-0.011*	-0.013	0.004
P-value	*0.806*	*0.617*	*0.732*	*0.790*	*0.778*	*0.099*	*0.105*	*0.662*

Notes: Based on propensity scores matching with 246 treated GPs. The size of the control group is 594 before 2012 421 in 2012, 325 in 2013 and 317 in 2014 and 2015.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

C. Potential Channels

There are several potential causal channels for the big reduction in hospitalization that we find. We have estimated the effect of the program on a series of intermediate outcomes which might lead to a reduction in hospitalization including number of GP visits, number of preventive visits, number

HbA1c blood tests, prescribed diabetes drugs (only observed until 2013), number of e-mail consultations, number of telephone consultations and number of influenza vaccinations, and Charlson index. We do not find any significant effects on any of those channels except for the last three. Both the number of influenza vaccinations and e-mail consultations are positively affected by the DMP in 2014 while we do find a reduction in the number of telephone consultations starting already from 2012 up to 2015 as well as on the Charlson index in 2014.

Overall our results are consistent with the idea that the DMP helped GPs being more systematic in the way they handled chronic patients which in turns lead to an improvement in their health status. This is in line with Lippert et al. 2014 who assess the DMP in a qualitative study based on several interviews and concluded that GPs used the information provided by the new system to improve "*…administration of a regular disease control schedule for patients with chronic disease and routine monitoring of outcomes for purposes of resource prioritization and medication management*" (M. L. Lippert et al., 2014).

D. Robustness checks

We have undertaken a range of robustness checks. First, we have estimated the effects in a population where we do not restrict the diabetic patients to be with the same GP in the period 2011-2015, which was performed to test whether treated GPs , because of the use of EHR, start identifying more diabetic patients and/or diabetic patients with less severe disease progression and hence less risk of hospitalization. If this is the case, the treatment effect may simply be due to the composition of the population and not a reduced risk of hospitalization per se. We have prevented this possibility by restricting the analysis to diabetic patients who were already diabetic before 2010, which we now relax. Table 3 (panel C) shows the estimated treatment effect with the unrestricted diabetes population. The estimated treatment effects are very comparable to the restricted population both in magnitude and significance. Hence, it is very unlikely that our results

are merely driven by composition bias.

We also estimate our effects using difference-in-differences where we condition on the same covariates as in the PSM, excluding the pre-treatment outcomes. Table 3 (panel D) shows that the treatment effects are comparable but a bit higher than the one of our main specification – especially for the second year after enrolment.

To assess the sensitivity of our results to different definitions of treatment intensity more explicitly, we estimate the effect with treatments defined replacing the 70% threshold with 60% and 80% respectively. Table 3 (panel A and B) shows that the effects are in general very similar.

As a final robustness check we aggregate our outcomes over the post-treatment years 2012-2014 or 2012-2015 and compare with the aggregated outcomes of four (2008-2010) or five pre-treatment years (2007-2010). As hospitalizations do not occur in every calendar year for most patients this should give a good indication of the total effect of the DMP over the entire post treatment period. Table 4 show that the effect is highly significant for all outcomes and close to three times higher than the results from the main specification in table 2.

TABLE 3: ATET ESTIMATE – ROBUSTNESS CHECKS

	Falsification				ATET			
	2008	2009	2010	2011	2012	2013	2014	2015
PANEL A: PSM 60% THRESHOLD								
Hospitalizations	-0.008	0.012	-0.013	-0.012	-0.014	-0.033*	-0.041*	-0.015
P-value	*0.346*	*0.235*	*0.250*	*0.324*	*0.343*	*0.065*	*0.053*	*0.435*
ACSC hospitalizations	-0.004	0.000	-0.005	-0.001	-0.002	-0.012*	-0.009	-0.006
P-value	*0.358*	*0.968*	*0.288*	*0.749*	*0.651*	*0.057*	*0.173*	*0.398*
Diabetes hospitalizations	0.001	-0.005	-0.007	-0.003	-0.004	-0.012*	-0.014*	-0.001
P-value	*0.787*	*0.371*	*0.194*	*0.536*	*0.474*	*0.062*	*0.090*	*0.911*
PANEL B: PSM 80% THRESHOLD								
Hospitalizations	-0.008	-0.002	0.001	-0.020	-0.013	-0.032	-0.055**	-0.016
P-value	*0.393*	*0.851*	*0.928*	*0.101*	*0.395*	*0.108*	*0.020*	*0.438*
ACSC hospitalizations	-0.003	0.000	-0.001	-0.002	-0.001	-0.012*	-0.013*	-0.007
P-value	*0.562*	*0.953*	*0.914*	*0.611*	*0.812*	*0.092*	*0.076*	*0.366*
Diabetes hospitalizations	0.000	-0.003	-0.001	-0.009	0.000	-0.010	-0.015	0.001
P-value	*0.931*	*0.522*	*0.793*	*0.122*	*0.994*	*0.242*	*0.105*	*0.907*
PANEL C: PSM UNRESTRICTED POPULATION								
Hospitalizations	-0.005	0.002	-0.007	-0.005	-0.016	-0.034*	-0.044**	-0.005
P-value	*0.522*	*0.844*	*0.547*	*0.659*	*0.283*	*0.061*	*0.045*	*0.820*
ACSC hospitalizations	-0.006	0.003	-0.005	0.004	-0.003	-0.011*	-0.012*	-0.006
P-value	*0.157*	*0.496*	*0.365*	*0.380*	*0.594*	*0.076*	*0.070*	*0.499*
Diabetes hospitalizations	-0.001	-0.002	-0.002	0.001	-0.002	-0.012*	-0.013	0.001
P-value	*0.806*	*0.617*	*0.732*	*0.817*	*0.719*	*0.082*	*0.112*	*0.910*
PANEL D: DID								
Hospitalizations	0.000	0.014	-0.014	-0.014	-0.007	-0.039**	-0.054**	-0.019
P-value	*0.986*	*0.202*	*0.195*	*0.206*	*0.630*	*0.018*	*0.005*	*0.294*
ACSC hospitalizations	0.001	0.008	-0.004	0.003	-0.004	-0.012*	-0.015**	-0.007
P-value	*0.749*	*0.110*	*0.407*	*0.483*	*0.553*	*0.081*	*0.038*	*0.402*
Diabetes hospitalizations	0.001	-0.006	0.000	-0.003	-0.006	-0.012*	-0.017**	0.000
P-value	0.906	0.271	0.990	0.569	0.305	0.073	0.031	0.976

Notes: The size of the treatment cohort reduces to 182 when the threshold is increased to 80% and increases to 273 when the threshold is reduced to 60%. The size of the control group is 594 before 2012. 421 in 2012. 325 in 2013 and 317 in 2014 and 2015.
PSM: Propensity score matching. DiD: Difference in difference.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE 4: ATET WITH AGGREGATED OUTCOME PERIODS

	2012-2014	2012-2015
Hospitalization	-0.127**	-0.111**
P-value	*0.035*	*0.011*
ACSC Hospitalizations	-0.043**	-0.027**
P-value	*0.018*	*0.039*
Diabetes hospitalizations	-0.048**	-0.042**
P-value	*0.047*	*0.021*

Notes: Based on the 2012 treatment cohort consisting of 246 GPs. The size of the control group is 421 before 2012, 325 in 2013 and 317 in 2014 and 2015. PSM: Propensity score matching. Outcomes has been aggregated over the periods 2012-2014 and 2012-2015 and compared with aggregated pre-treatment years (2008-10) and (2007-2010) respectively.

VII. Discussion

We have estimated the effect of GP participation in an EHR program introduced in Denmark with the aim of improving primary care for diabetic patients. Our results show that the introduction of the EHR reduces hospitalization, and diabetes hospitalizations with more than 10% and ACSC hospitalizations by almost 16%. Our findings are robust to several sensitivity checks and different models. The reduction in hospitalizations is comparable or higher than the previous results in the related literature – even for DMP's with much stronger incentives as, for example, pay for performance or public feedback reports.

APPENDIX

TABLE A1: Probit model – probability of treatment (Hospitalizations)

Covariate	coef	p
Planned preventive visits	0.07	0.26
GP visits	-0.03	0.38
Influenza vaccinations	0.69	0.13
Blood tests/HbA1c	-0.15**	0.03
Email consultations	0.07	0.29
listsize2010	0.00**	0.04
charlson_index2010	-1.49***	0
Diabetes drugs	0	0.41
Telephone consultations	-0.06**	0.01
DADI 2	1.18	0.17
DADI_12010	-2.58	0.28
DADI 2	1.22	0.5
DADI 4	2.59**	0.01
DADI 8	-0.47	0.57
DADI 6	-1.74**	0.03
DADI 7	-1.50**	0.04
Region 2	-0.16	0.42
Region 3	0.26	0.18
Region 4	0.47**	0.04
Region 5	-0.15	0.53
Hospitalizations2010	0.17	0.69
Hospitalizations2009	0.73	0.13
Hospitalizations2008	-0.23	0.66
Hospitalizations2007	-0.89*	0.1
Hospitalizations2006	0.22	0.69
Hospitalizations2005	0.29	0.61
Hospitalizations2004	0.53	0.34
Constant	0.31	0.58

TABLE A2: Probit model – probability of treatment (Diabetes hospitalizations)

Covariate	coef	p
Planned preventive visits	0.09	0.17
GP visits	-0.02	0.52
Influenza vaccinations	0.69	0.13
Blood tests/HbA1c	-0.16**	0.02
Email consultations	0.06	0.37
listsize2010	0.00**	0.04
charlson_index2010	-1.28***	0
Diabetes drugs	0	0.49
Telephone consultations	-0.06**	0.01
DADI 2	1.27	0.14
DADI_12010	-2.21	0.34
DADI 2	1.2	0.5
DADI 4	2.72***	0.01
DADI 8	-0.33	0.68
DADI 6	-1.93**	0.02
DADI 7	-1.19	0.1
Region 2	-0.13	0.49
Region 3	0.3	0.13
Region 4	0.54**	0.02
Region 5	-0.04	0.85
Diabetes hospitalizations2010	0.18	0.86
Diabetes hospitalizations2009	-0.82	0.45
Diabetes hospitalizations2008	0.1	0.92
Diabetes hospitalizations2007	-0.33	0.79
Diabetes hospitalizations2006	-0.42	0.74
Diabetes hospitalizations2005	-0.89	0.54
Diabetes hospitalizations2004	1.24	0.39
Constant	0.27	0.63

TABLE A3: Probit model – probability of treatment (ACSC hospitalizations)

Covariate	coef	p
Planned preventive visits	0.08	0.2
GP visits	-0.02	0.56
Influenza vaccinations	0.68	0.14
Blood tests/HbA1c	-0.17**	0.02
Email consultations	0.06	0.38
listsize2010	0.00**	0.03
charlson_index2010	-1.26***	0
Diabetes drugs	0	0.47
Telephone consultations	-0.06**	0.01
DADI 2	1.2	0.16
DADI_12010	-1.87	0.43
DADI 2	1.23	0.49
DADI 4	2.39**	0.02
DADI 8	-0.33	0.68
DADI 6	-1.87**	0.02
DADI 7	-1.40*	0.06
Region 2	-0.14	0.49
Region 3	0.28	0.15
Region 4	0.55**	0.02
Region 5	-0.04	0.84
ACSC Hospitalizations2010	0	1
ACSC Hospitalizations2009	0.07	0.94
ACSC Hospitalizations2008	-0.99	0.35
ACSC Hospitalizations2007	-1.81	0.12
ACSC Hospitalizations2006	0.92	0.33
ACSC Hospitalizations2005	0.2	0.85
ACSC Hospitalizations2004	-0.11	0.93
Constant	0.39	0.49

TABLE A4: Bias reduction - Hospitalizations

Variable	Unmatched Matched	Treated	control	%bias	%reduct	t	p>t
Planned preventive visits	U	1.394	1.237	19.500		2.390	0.017
	M	1.400	1.345	6.800	65.100	0.770	0.439
GP visits	U	6.952	6.778	11.300		1.400	0.163
	M	6.973	6.977	-0.300	97.500	-0.030	0.974
Influenza vaccinations	U	0.303	0.309	-5.500		-0.680	0.494
	M	0.305	0.304	0.700	87.900	0.070	0.940
Blood tests/HbA1c	U	1.259	1.335	-8.800		-1.100	0.273
	M	1.275	1.272	0.300	96.600	0.030	0.972
Email consultations	U	0.824	0.728	12.300		1.490	0.136
	M	0.810	0.820	-1.300	89.200	-0.150	0.884
Listsize2010	U	2353.200	1932.900	33.800		4.250	0.000
	M	2355.900	2324.300	2.500	92.500	0.250	0.801
Charlson_index2010	U	0.505	0.560	-29.300		-3.600	0.000
	M	0.504	0.506	-1.000	96.600	-0.120	0.901
Diabetes drugs	U	984.170	1051.900	-16.600		-2.060	0.039
	M	984.150	976.550	1.900	88.800	0.210	0.831
Telephone consultations	U	5.011	5.618	-29.300		-3.570	0.000
	M	5.060	5.195	-6.600	77.600	-0.770	0.439
DADI 2	U	0.278	0.269	12.200		1.520	0.130
	M	0.277	0.275	2.500	79.200	0.280	0.778
DADI_1	U	0.038	0.041	-9.600		-1.180	0.239
	M	0.038	0.037	1.200	87.600	0.140	0.891
DADI 2	U	0.085	0.096	-16.800		-2.070	0.039
	M	0.085	0.085	0.500	96.900	0.060	0.952
DADI 4	U	0.118	0.101	28.200		3.550	0.000
	M	0.116	0.114	3.700	86.800	0.410	0.682
DADI 8	U	0.256	0.277	-22.900		-2.810	0.005
	M	0.258	0.256	2.200	90.300	0.250	0.800
DADI 6	U	0.092	0.110	-16.800		-2.090	0.037

	M	0.092	0.093	-0.700	96.000	-0.080	0.936
DADI 7	U	0.314	0.337	-30.100		-3.700	0.000
	M	0.314	0.317	-3.900	87.100	-0.470	0.641
Region 2	U	0.207	0.249	-10.000		-1.240	0.216
	M	0.212	0.215	-0.900	90.800	-0.100	0.918
Region 3	U	0.256	0.211	10.600		1.330	0.185
	M	0.257	0.257	0.000	99.800	0.000	0.998
Region 4	U	0.321	0.297	5.200		0.650	0.513
	M	0.311	0.312	-0.100	98.400	-0.010	0.992
Region 5	U	0.118	0.157	-11.300		-1.390	0.166
	M	0.120	0.116	1.200	89.300	0.140	0.888
Hospitalizations2010	U	0.397	0.419	-15.100		-1.850	0.065
	M	0.398	0.397	0.700	95.100	0.090	0.929
Hospitalizations2009	U	0.372	0.365	5.700		0.690	0.491
	M	0.370	0.366	3.600	36.400	0.430	0.671
Hospitalizations2008	U	0.320	0.326	-5.500		-0.670	0.502
	M	0.319	0.322	-3.100	43.800	-0.370	0.715
Hospitalizations2007	U	0.301	0.318	-15.700		-1.920	0.055
	M	0.303	0.304	-0.600	96.200	-0.070	0.943
Hospitalizations2006	U	0.299	0.301	-1.200		-0.150	0.880
	M	0.298	0.295	2.600	-116.000	0.310	0.760
Hospitalizations2005	U	0.276	0.278	-1.900		-0.230	0.815
	M	0.273	0.273	0.300	85.100	0.030	0.975
Hospitalizations2004	U	0.269	0.263	5.800		0.710	0.476
	M	0.269	0.263	6.100	-6.300	0.760	0.448

TABLE A5: Bias reduction – Diabetes hospitalizations

Variable	Unmatched Matched	Treated	control	%bias	%reduct	t	p>t
Planned preventive visits	U	1.394	1.237	19.500		2.390	0.017
	M	1.394	1.378	2.000	89.900	0.230	0.820
GP visits	U	6.952	6.778	11.300		1.400	0.163
	M	6.952	7.025	-4.700	58.400	-0.550	0.581
Influenza vaccinations	U	0.303	0.309	-5.500		-0.680	0.494
	M	0.303	0.306	-2.900	46.800	-0.330	0.741
Blood tests/HbA1c	U	1.259	1.335	-8.800		-1.100	0.273
	M	1.259	1.272	-1.500	83.400	-0.170	0.865
Email consultations	U	0.824	0.728	12.300		1.490	0.136
	M	0.824	0.830	-0.800	93.700	-0.090	0.932
Listsize2010	U	2353.200	1932.900	33.800		4.250	0.000
	M	2353.200	2260.900	7.400	78.000	0.770	0.440
Charlson_index2010	U	0.505	0.560	-29.300		-3.600	0.000
	M	0.505	0.509	-2.100	92.700	-0.270	0.789
Diabetes drugs	U	984.170	1051.900	-16.600		-2.060	0.039
	M	984.170	981.510	0.700	96.100	0.080	0.940
Telephone consultations	U	5.011	5.618	-29.300		-3.570	0.000
	M	5.011	5.041	-1.500	95.100	-0.170	0.865
DADI 2	U	0.278	0.269	12.200		1.520	0.130
	M	0.278	0.276	2.100	82.700	0.230	0.817
DADI_1	U	0.038	0.041	-9.600		-1.180	0.239
	M	0.038	0.037	1.300	86.600	0.150	0.880
DADI 2	U	0.085	0.096	-16.800		-2.070	0.039
	M	0.085	0.083	3.700	78.000	0.440	0.662
DADI 4	U	0.118	0.101	28.200		3.550	0.000
	M	0.118	0.114	5.900	79.100	0.660	0.509
DADI 8	U	0.256	0.277	-22.900		-2.810	0.005
	M	0.256	0.250	6.300	72.600	0.710	0.481
DADI 6	U	0.092	0.110	-16.800		-2.090	0.037

	M	0.092	0.095	-2.600	84.600	-0.310	0.754
DADI 7	U	0.314	0.337	-30.100		-3.700	0.000
	M	0.314	0.311	3.900	87.100	0.460	0.643
Region 2	U	0.207	0.249	-10.000		-1.240	0.216
	M	0.207	0.208	-0.100	98.700	-0.010	0.988
Region 3	U	0.256	0.211	10.600		1.330	0.185
	M	0.256	0.256	0.000	100.000	0.000	1.000
Region 4	U	0.321	0.297	5.200		0.650	0.513
	M	0.321	0.321	0.100	97.700	0.010	0.990
Region 5	U	0.118	0.157	-11.300		-1.390	0.166
	M	0.118	0.118	0.000	100.000	0.000	1.000
Diabetes hospitalizations2010	U	0.099	0.113	-20.500		-2.470	0.014
	M	0.099	0.097	3.300	83.700	0.440	0.660
Diabetes hospitalizations2009	U	0.088	0.095	-12.200		-1.470	0.142
	M	0.088	0.085	6.500	46.300	0.780	0.435
Diabetes hospitalizations2008	U	0.077	0.085	-12.800		-1.610	0.108
	M	0.077	0.074	4.800	62.200	0.580	0.564
Diabetes hospitalizations2007	U	0.067	0.074	-14.800		-1.800	0.072
	M	0.067	0.063	8.200	44.500	1.010	0.314
Diabetes hospitalizations2006	U	0.062	0.071	-19.500		-2.370	0.018
	M	0.062	0.061	1.400	92.700	0.180	0.859
Diabetes hospitalizations2005	U	0.054	0.063	-20.100		-2.450	0.015
	M	0.054	0.052	4.900	75.900	0.590	0.557
Diabetes hospitalizations2004	U	0.053	0.055	-6.000		-0.740	0.458
	M	0.053	0.050	7.800	-30.200	0.960	0.339

TABLE A6: Bias reduction – ACSC hospitalizations

Variable	Unmatched Matched	Treated	control	%bias	%reduct	t	p>t
Planned preventive visits	U	1.394	1.237	19.500		2.390	0.017
	M	1.394	1.378	2.000	89.900	0.230	0.820
GP visits	U	6.952	6.778	11.300		1.400	0.163
	M	6.952	7.025	-4.700	58.400	-0.550	0.581
Influenza vaccinations	U	0.303	0.309	-5.500		-0.680	0.494
	M	0.303	0.306	-2.900	46.800	-0.330	0.741
Blood tests/HbA1c	U	1.259	1.335	-8.800		-1.100	0.273
	M	1.259	1.272	-1.500	83.400	-0.170	0.865
Email consultations	U	0.824	0.728	12.300		1.490	0.136
	M	0.824	0.830	-0.800	93.700	-0.090	0.932
Listsize2010	U	2353.200	1932.900	33.800		4.250	0.000
	M	2353.200	2260.900	7.400	78.000	0.770	0.440
Charlson_index2010	U	0.505	0.560	-29.300		-3.600	0.000
	M	0.505	0.509	-2.100	92.700	-0.270	0.789
Diabetes drugs	U	984.170	1051.900	-16.600		-2.060	0.039
	M	984.170	981.510	0.700	96.100	0.080	0.940
Telephone consultations	U	5.011	5.618	-29.300		-3.570	0.000
	M	5.011	5.041	-1.500	95.100	-0.170	0.865
DADI 2	U	0.278	0.269	12.200		1.520	0.130
	M	0.278	0.276	2.100	82.700	0.230	0.817
DADI_1	U	0.038	0.041	-9.600		-1.180	0.239
	M	0.038	0.037	1.300	86.600	0.150	0.880
DADI 2	U	0.085	0.096	-16.800		-2.070	0.039
	M	0.085	0.083	3.700	78.000	0.440	0.662
DADI 4	U	0.118	0.101	28.200		3.550	0.000
	M	0.118	0.114	5.900	79.100	0.660	0.509
DADI 8	U	0.256	0.277	-22.900		-2.810	0.005
	M	0.256	0.250	6.300	72.600	0.710	0.481
DADI 6	U	0.092	0.110	-16.800		-2.090	0.037

	M	0.092	0.095	-2.600	84.600	-0.310	0.754
DADI 7	U	0.314	0.337	-30.100		-3.700	0.000
	M	0.314	0.311	3.900	87.100	0.460	0.643
Region 2	U	0.207	0.249	-10.000		-1.240	0.216
	M	0.207	0.208	-0.100	98.700	-0.010	0.988
Region 3	U	0.256	0.211	10.600		1.330	0.185
	M	0.256	0.256	0.000	100.000	0.000	1.000
Region 4	U	0.321	0.297	5.200		0.650	0.513
	M	0.321	0.321	0.100	97.700	0.010	0.990
Region 5	U	0.118	0.157	-11.300		-1.390	0.166
	M	0.118	0.118	0.000	100.000	0.000	1.000
ACSC hospitalizations2010	U	0.099	0.113	-20.500		-2.470	0.014
	M	0.099	0.097	3.300	83.700	0.440	0.660
ACSC hospitalizations2009	U	0.088	0.095	-12.200		-1.470	0.142
	M	0.088	0.085	6.500	46.300	0.780	0.435
ACSC hospitalizations2008	U	0.077	0.085	-12.800		-1.610	0.108
	M	0.077	0.074	4.800	62.200	0.580	0.564
ACSC hospitalizations2007	U	0.067	0.074	-14.800		-1.800	0.072
	M	0.067	0.063	8.200	44.500	1.010	0.314
ACSC hospitalizations2006	U	0.062	0.071	-19.500		-2.370	0.018
	M	0.062	0.061	1.400	92.700	0.180	0.859
ACSC hospitalizations2005	U	0.054	0.063	-20.100		-2.450	0.015
	M	0.054	0.052	4.900	75.900	0.590	0.557
ACSC hospitalizations2004	U	0.053	0.055	-6.000		-0.740	0.458
	M	0.053	0.050	7.800	-30.200	0.960	0.339

REFERENCES

(AHQR), Agency for Healthcare Quality and Research. 2001. "Ahrq Quality Indicators—Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions.," Rockville, MD: Agency for Healthcare Research and Quality,

Abadie, A. and G.W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica*, 74(1), 32.

Berger, L.; H. Bleichrodt and L. Eeckhoudt. 2013. "Treatment Decisions under Ambiguity." *J Health Econ*, 32(3), 559-69.

Chabé-Ferret, S. 2015. "Analysis of the Bias of Matching and Difference-in-Difference under Alternative Earnings and Selection Processes." *Journal of Econometrics*, 185(1), 13.

Chalkley, M. and J. M. Malcomson. 1998. "Contracting for Health Services When Patient Demand Does Not Reflect Quality." *J Health Econ*, 17(1), 1-19.

Dusheiko, M.; T. Doran; H. Gravelle; C. Fullwood and M. Roland. 2011. "Does Higher Quality of Diabetes Management in Family Practice Reduce Unplanned Hospital Admissions?" *Health Serv Res*, 46(1 Pt 1), 27-46.

Eijkenaar, F.; M. Emmert; M. Scheppach and O. Schoffski. 2013. "Effects of Pay for Performance in Health Care: A Systematic Review of Systematic Reviews." *Health Policy*, 110(2-3), 115-30.

Ellis, R. P. and T. G. McGuire. 1986. "Provider Behavior under Prospective Reimbursement. Cost Sharing and Supply." *J Health Econ*, 5(2), 129-51.

Ellrodt, G.; D. J. Cook; J. Lee; M. Cho; D. Hunt and S. Weingarten. 1997. "Evidence-Based Disease Management." *JAMA*, 278(20), 1687-92.

Epstein, A. J. 2006. "Do Cardiac Surgery Report Cards Reduce Mortality? Assessing the Evidence." *Med Care Res Rev*, 63(4), 403-26.

Gans, D.; J. Kralewski; T. Hammons and B. Dowd. 2005. "Medical Groups' Adoption of Electronic Health Records and Information Systems." *Health Aff (Millwood)*, 24(5), 1323-33.

Gaynor, M.; Pauly, MV. 1990. "Compensation and Productive Efficiency in Partnerships: Evidence from Medical Groups Practice." *Journal of Political Economy*, 98(3), 29.

Han, W.; R. Sharman; A. Heider; N. Maloney; M. Yang and R. Singh. 2016. "Impact of Electronic Diabetes Registry 'Meaningful Use' on Quality of Care and Hospital Utilization." *J Am Med Inform Assoc*, 23(2), 242-7.

Harrison, M. J.; M. Dusheiko; M. Sutton; H. Gravelle; T. Doran and M. Roland. 2014. "Effect of a National Primary Care Pay for Performance Scheme on Emergency Hospital Admissions for Ambulatory Care Sensitive Conditions: Controlled Longitudinal Study." *BMJ*, 349, g6423.

Huber, Martin; Michael Lechner and Andreas Steinmayr. 2015. "Radius Matching on the Propensity Score with Bias Adjustment: Tuning Parameters and Finite Sample Behaviour." *Empirical Economics*, 49(1), 1-31.

Huber, Martin; Michael Lechner and Conny Wunsch. 2013. "The Performance of Estimators Based on the Propensity Score." *Journal of Econometrics*, 175(1), 1-21.

Iezzi, E.; M. Lippi Bruni and C. Ugolini. 2014. "The Role of Gp's Compensation Schemes in Diabetes Care: Evidence from Panel Data." *J Health Econ*, 34, 104-20.

Imbens, G. and J. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1), 5-85.

Iversen, T. 2004. "The Effects of a Patient Shortage on General Practitioners' Future Income and List of Patients." *J Health Econ*, 23(4), 673-94.

Kasteridis, P.; A. Street; M. Dolman; L. Gallier; K. Hudson; J. Martin and I. Wyer. 2015. "Who Would Most Benefit from Improved Integrated Care? Implementing an Analytical Strategy in South Somerset." *Int J Integr Care*, 15, e001.

Ketcham, J. D.; L. C. Baker and D. MacIsaac. 2007. "Physician Practice Size and Variations in Treatments and Outcomes: Evidence from Medicare Patients with Ami." *Health Aff (Millwood)*, 26(1), 195-205.

Knight, K.; E. Badamgarav; J. M. Henning; V. Hasselblad; A. D. Gano, Jr.; J. J. Ofman and S. R. Weingarten. 2005. "A Systematic Review of Diabetes Disease Management Programs." *Am J Manag Care*, 11(4), 242-50.

Kolstad, J.T. 2013. "Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards." *American Economic Review*, 103(7), 2875-910.

Lechner, M. 2015. "Treatment Effects and Panel Data," B. Baltagi, *The Oxford Handbook of Panel Data*. Oxford University Press,

Lippert, L. 2014. *What (Quality)Development in General Practice?* Copenhagen: University of Copenhagen, Health Faculty.

Lippert, M. L.; M. B. Kousgaard and L. Bjerrum. 2014. "General Practitioners Uses and Perceptions of Voluntary Electronic Feedback on Treatment Outcomes - a Qualitative Study." *BMC Fam Pract*, 15, 193.

Lyttkens, C.H.; Christiansen, T.; Häkkinen, U.; Karboe, O.; Welander, A. . 2016. "The Core of the Nordic Health Care System Is Not Empty." *Nordic Journal of Health Economics*, 4(1), 20.

Ma, A; Mak, H. Forthcoming. "Incentives in Health Care Payment Systems." *Oxford Research Encyclopedia of Economics and Finance*.

McGinn, C. A.; S. Grenier; J. Duplantie; N. Shaw; C. Sicotte; L. Mathieu; Y. Leduc; F. Legare and M. P. Gagnon. 2011. "Comparison of User Groups' Perspectives of Barriers and Facilitators to Implementing Electronic Health Records: A Systematic Review." *BMC Med*, 9, 46.

Olsen, K. R. 2012. "Patient Complexity and Gps' Income under Mixed Remuneration." *Health Econ*, 21(6), 619-32.

Olsen, K.R. ; A. Anell; U. Häkkinen; T. Iversen; T. Ólafsdóttir and M. Sutton. 2016. "General Practice in the Nordic Countries." *Nordic Journal of Health Economics*, 4(1), 56-67.

Patel, V.; M. E. Reed and R. W. Grant. 2015. "Electronic Health Records and the Evolution of Diabetes Care: A Narrative Review." *J Diabetes Sci Technol*, 9(3), 676-80.

Reed, M.; J. Huang; R. Brand; I. Graetz; R. Neugebauer; B. Fireman; M. Jaffe; D. W. Ballard and J. Hsu. 2013. "Implementation of an Outpatient Electronic Health Record and Emergency Department Visits, Hospitalizations, and Office Visits among Patients with Diabetes." *JAMA*, 310(10), 1060-5.

Rosenbaum, P. and D. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70(1), 11.

Rubin, D. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology*, 66, 13.

Schrijvers, G. 2009. "Disease Management: A Proposal for a New Definition." *Int J Integr Care*, 9, e06.

Schroll, H.; R. D. Christensen; J. L. Thomsen; M. Andersen; S. Friborg and J. Sondergaard. 2012. "The Danish Model for Improvement of Diabetes Care in General Practice: Impact of Automated Collection and Feedback of Patient Data." *Int J Family Med*, 2012, 208123.

Scott, A. 2000. "Economics of General Practice," A. J. N. Culyer, J.P., *Handbook of Health Economics.* Elsevier Science B.V.,

Steinbrook, R. 2006. "Public Report Cards--Cardiac Surgery and Beyond." *N Engl J Med*, 355(18), 1847-9.

Ulstrup, F. 2012. "Datafangst Giver Overblik Og Højere Kvalitet." *Danske bioanalytikere*, 2012(1), 12-15.

Wang, Y.; C. A. O'Donnell; D. F. Mackay and G. C. Watt. 2006. "Practice Size and Quality Attainment under the New Gms Contract: A Cross-Sectional Analysis." *Br J Gen Pract*, 56(532), 830-5.