


# The Effect of Financial Incentives on Quality Measures in the Treatment of Diabetes Mellitus: a Randomized Controlled Trial



Rahel Meier, Dr.<sup>1,2</sup> , Corinne Chmiel, Dr.<sup>1,2</sup>, Fabio Valeri<sup>1,2</sup>, Leander Muheim, Dr.<sup>1,2</sup>, Oliver Senn, Prof.<sup>1,2</sup>, and Thomas Rosemann, Prof.<sup>1,2</sup>

<sup>1</sup>Institute of Primary Care, University of Zurich, Pestalozzistrasse 24, Zurich, Switzerland; <sup>2</sup>University Hospital Zurich, Zürich, Switzerland.

**BACKGROUND:** Financial incentives are often used to improve quality of care in chronic care patients. However, the evidence concerning the effect of financial incentives is still inconclusive.

**OBJECTIVE:** To test the effect of financial incentives on quality measures (QMs) in the treatment of patients with diabetes mellitus in primary care. We incentivized a clinical QM and a process QM to test the effect of financial incentives on different types of QMs and to investigate the spill-over effect on non-incentivized QMs.

**DESIGN/PARTICIPANTS:** Parallel cluster randomized controlled trial based on electronic medical records database involving Swiss general practitioners (GPs). Practices were randomly allocated.

**INTERVENTION:** All participants received a bimonthly feedback report. The intervention group additionally received potential financial incentives on GP level depending on their performance.

**MAIN MEASURES:** Between-group differences in proportions of patients fulfilling incentivized QM (process QM of annual HbA1c measurement and clinical QM of blood pressure level below 140/95 mmHg) after 12 months.

**KEY RESULTS:** Seventy-one GPs (median age 52 years, 72% male) from 43 different practices and subsequently 3838 patients with diabetes mellitus (median age 70 years, 57% male) were included. Proportions of patients with annual HbA1c measurements remained unchanged (intervention group decreased from 79.0 to 78.3%, control group from 81.5 to 81.0%, OR 1.09, 95% CI 0.90–1.32,  $p = 0.39$ ). Proportions of patients with blood pressure below 140/95 improved from 49.9 to 52.5% in the intervention group and decreased from 51.2 to 49.0% in the control group (OR 1.16, 95% CI 0.99–1.36,  $p = 0.06$ ). Proportions of non-incentivized process QMs increased significantly in the intervention group.

**CONCLUSION:** GP level financial incentives did not result in more frequent HbA1c measurements or in improved blood pressure control. Interestingly, we could confirm a spill-over effect on non-incentivized process QMs. Yet, the mechanism of spill-over effects of financial incentives is largely unclear.

**TRIAL REGISTRATION:** ISRCTN13305645

**KEY WORDS:** quality of care; improvement; financial incentives; diabetes mellitus; primary care; spill-over effect.

J Gen Intern Med 37(3):556–64

DOI: 10.1007/s11606-021-06714-8

© The Author(s) 2021

## BACKGROUND

Improving quality of care in chronic care patients, especially in patients with diabetes mellitus, has been of major priority around the world.<sup>1</sup> Different strategies, such as audit and feedback, clinical education, clinical reminders, or financial incentives, have been used, to try to close the gap between actual and optimal care.<sup>2</sup> Financial incentives, also called pay for performance (P4P) strategies, have been studied in different settings. However, evidence concerning the effect on quality of care is still inconclusive.<sup>3–6</sup>

Many studies investigating the effect of P4P were conducted in the UK, as they pioneered a P4P strategy with its Quality and Outcomes Framework (QOF). Since it was a nationwide pre/post real-life setting without a control group, its effect remains controversial.<sup>7,8</sup> Randomized controlled trials (RCTs) conducted in the USA or Taiwan showed only modest improvements from P4P interventions,<sup>9–14</sup> whereas evidence from RCTs in Europe is lacking. In Switzerland, no data on the P4P approach exists and the use of quality measures (QMs), especially in primary care, has been marginal. In Switzerland, general practitioners (GPs) are paid on a fee-for-service basis, based on tariffs contractually agreed between insurers and providers or set by the government.<sup>15</sup>

With the current parallel cluster RCT, we aimed to test whether financial incentives are effective in increasing proportions of patients with diabetes mellitus meeting a QM. Of six diabetes-specific QMs, we incentivized a process QM and a clinical QM to test the effect of incentives on different types of QMs and to investigate spill-over effects on non-incentivized QMs.

## METHODS

### Trial Design

We conducted a parallel cluster RCT with GPs participating in the FIRE network (family medicine International Classification

Received November 23, 2020

Accepted March 9, 2021

Published online April 26, 2021

of Primary Care (ICPC) research using electronic medical records (EMRs)).<sup>16</sup> GPs participating in the FIRE network contribute anonymized patient data to the steadily growing database containing the following data: administrative information, vital signs, laboratory values, medication data, and diagnostic codes, coded in the ICPC-2 classification scheme.<sup>17</sup>

According to the local ethics committee of the canton of Zurich, the project does not fall under the scope of the law on human research and therefore, no ethical consent was necessary (BASEC-Nr. Req-2017-00797). The trial was registered in the ISRCTN registry (identifier: ISRCTN13305645) and the protocol has been published.<sup>18</sup>

## Participants

FIRE GPs had additionally to fulfill the following data availability and data quality criteria to be eligible for the RCT: (a) continuous data delivery since January 2017, (b) delivering HbA1c and blood pressure values in more than 10% of their patients with diabetes mellitus, and (c) treating a minimum of five patients with diabetes mellitus. In June 2018, the eligible GPs received an invitation to participate in the RCT (Fig. 1).

From participating GPs, we included all patients diagnosed with diabetes mellitus if they were diagnosed at least 4 months before the baseline assessment and had at least one consultation during the last 12 months. During the observation period, patients with newly identified diabetes mellitus were included whereas patients not having a consultation within the last 12 months or reported dead dropped out. Patients with diabetes were identified if they met at least one of the following criteria: (1) patients with ICPC-2 codes T89 (insulin-dependent diabetes mellitus) and T90 (insulin-independent diabetes mellitus); (2) patients with antidiabetic medication according to the Anatomical Therapeutic Chemical (ATC) classification system (A10A, A10B, A10X).<sup>19</sup>

## Intervention

Both groups received a bimonthly diabetes feedback report containing information on their patients with diabetes mellitus (age, gender, and body mass index), the proportion of patients receiving at least one HbA1c measurement within the last 12 months, and the proportions of patients with blood pressure measurements and achieving the target blood pressure level. Furthermore, the report contained a key message addressing various issues in the treatment of patients with diabetes mellitus. No patient-specific feedback could be included due to anonymized patient data. An example of the report can be found in Additional File 1. The report was designed to inform GPs about their current achievements in primary outcomes and was used as a sham intervention for the control group. We informed the intervention group at the beginning of the observation period and in a reminder after 6 months about the incentive at GP level. We announced that they will receive an incentive of 75 Swiss francs per improved percentage point in the two QMs mentioned in the feedback report after the

observation period. The amount of 75 Swiss francs corresponds to an average net hourly revenue of a Swiss GP.<sup>20</sup> The control group was blinded for the incentives of the intervention group.

## Outcomes

Primary outcome (two independent primary endpoints): between-group differences in the proportions of patients fulfilling incentivized QMs after 12 months (Table 1).

Secondary outcome: between-group differences in the proportions of patients fulfilling non-incentivized QMs after 12 months (Table 1).

We assessed QM proportions at baseline—based on the 12 months prior to the study start—and every 2 months until the observational period ended 12 months after a GP was included in the study.

## Sample Size

We adjusted the sample size calculation reported in the study protocol, as investigations showed that the assumed baseline proportions have increased. We newly assumed an improvement for the process QM (HbA1c) from currently 80 to 90%. For the clinical QM (controlled blood pressure), we assumed an improvement from currently 60 to 70%. We chose a power of 80% and a type-I error of 5% and we assumed a median number of two GPs per practice and a median number of 30 patients with diabetes mellitus per practice. Furthermore, for all outcomes, we estimated an intraclass correlation at GP level and at patient level based on the FIRE database (for further details see, Additional File 2). Based on these assumptions and a sample size calculation taking the multilevel cluster setting into account, the number of GPs needed for the study remained at 70, hence unchanged to the study protocol.<sup>21</sup>

## Randomization

The randomization took place on practice level to avoid contamination (intervention and control within the same practice). The current proportion of the clinical QM, number of participating GPs per practice, GP network participation, and number of patients with diabetes mellitus were used to stratify randomization. We used the shiny balancer software for randomization<sup>22</sup> and generated 100 balanced allocation schemes. The settings for the shiny balancer software are displayed in Additional File 1, Table 1. We applied two additional constraints to the generated schemes: (1) the two largest practices were not in the same study group; (2) the difference of number of GPs between the control and intervention group was minimal. We selected one of the proposed schemes randomly.

## Sensitivity Analysis

For a sensitivity analysis, we determined proportions of patients fulfilling the QMs of GPs not willing to participate. The analysis was performed retrospectively for the same time

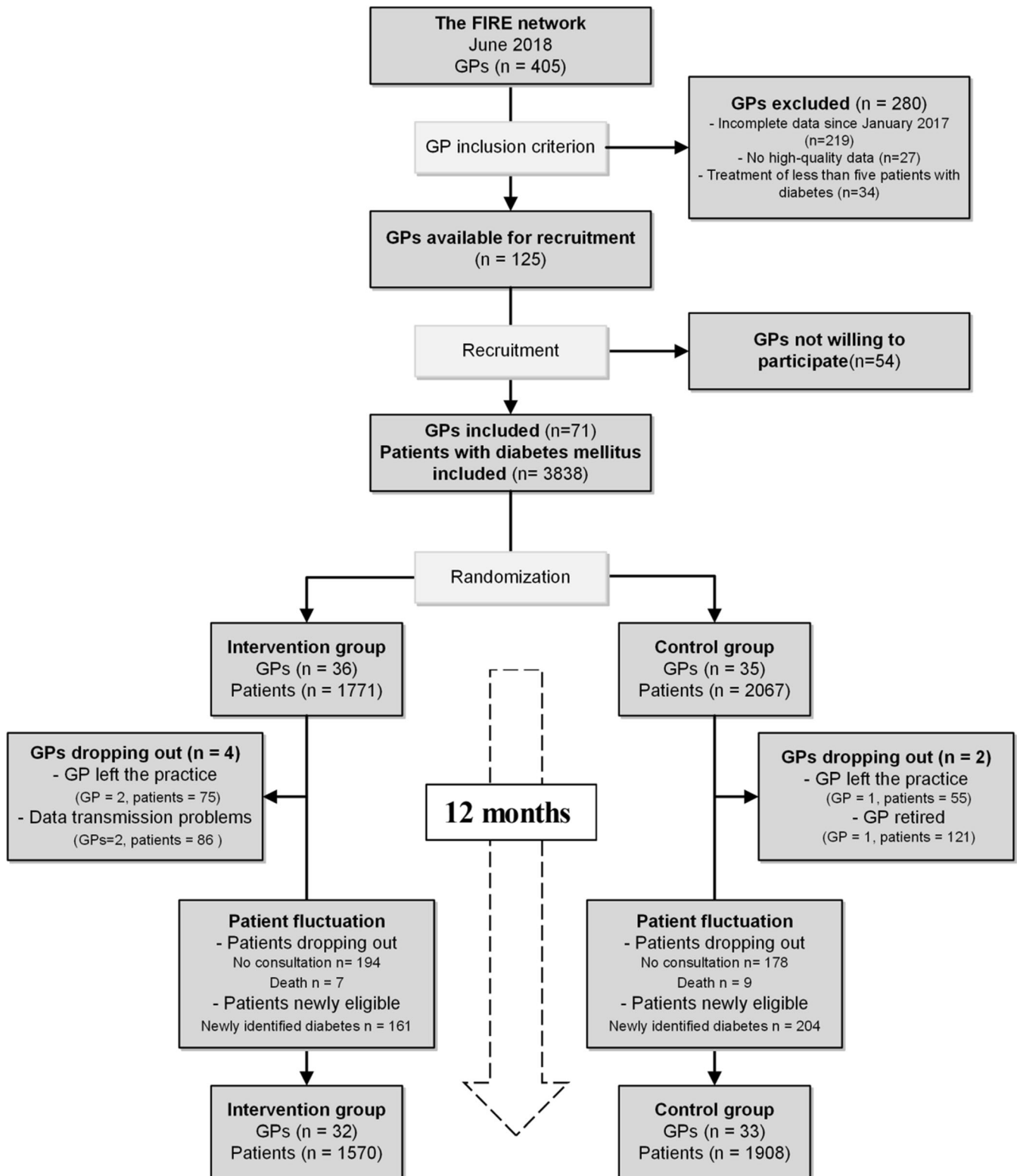


Figure 1 Flowchart of the study including dropouts of GPs and patient fluctuation. GP general practitioner, pat patient.

points as for the GPs in the study. With this analysis, we aimed to investigate whether some independent improvements in QMs have occurred simultaneously and if the intervention of solely receiving educational feedback reports had an effect.

## Statistical Methods

We reported categorical data as frequencies and percentages, and continuous variables as means and standard deviations (SD) or median and interquartile range (IQR), as appropriate. To assess the effect of financial incentives, we used

**Table 1 Quality Measures Used to Assess Primary and Secondary Outcomes**

Type	Measure	Description
Incentivized QMs		
Clinical QM	Blood pressure	Proportion of patients with diabetes mellitus with last blood pressure measurement < 140/85 mmHg in the preceding 12 months.
Process QM	HbA1c	Proportion of patients with diabetes mellitus with at least one measurement of HbA1c in the preceding 12 months.
Non-incentivized QMs		
Process QM	Blood pressure	Proportion of patients with diabetes mellitus with at least one blood pressure measurement in the preceding 12 months.
Clinical QM	HbA1c	Proportion of patients with diabetes mellitus with HbA1c levels < 7.5% in the preceding 12 months.
Process QM	Cholesterol	Proportion of patients with diabetes mellitus with at least one cholesterol measurement in the preceding 12 months.
Clinical QM	Cholesterol	Proportion of patients with diabetes mellitus with total cholesterol < 5 mmol/L in the preceding 12 months.

*QM quality measures*

hierarchical multivariable logistic regression models, with random variables of patients nested in GPs and GPs nested in practice. The QM fulfillment was the dependent variable, and independent variables were time and group allocation, for which we allowed interaction. Furthermore, we adjusted for age and gender of the GPs and the volume of patients with diabetes mellitus per GP. The same model was used for the sensitivity analysis, with the non-participating GPs as a third group allocation. To visualize trend and effect, we computed for each GP the percentage of patients fulfilling the QMs and aggregated these according to group allocation and time and computed the mean and the Wald 95% confidence interval.

**RESULTS****Study Population**

Of the 125 GPs identified as eligible and invited to participate, 71 gave consent to participate. Due to difficulties in recruitment, 61 GPs started in September 2018 in cohort 1, whereas 10 GPs were enrolled in a second cohort starting in November 2018. Randomization took place at the practice level (intervention group: 21 practices and 36 GPs, control group: 22 practices and 35 GPs). Subsequently, all their patients with diabetes mellitus were included, resulting in a starting population of 3838 patients (intervention group: 1771 patients, control group: 2067 patients). During the 12-month observation period, four GPs dropped out in the intervention group (161 patients) and two in the control group (176 patients). Reasons are outlined in Figure 1 and information on dropout date of GPs is available in Additional File 1, Table 2. Patient fluctuation due to changes in eligibility status (death, no consultation within 12 months, or newly identified patients with diabetes mellitus) was minimal and more than 85% of patients were observed during the entire observation period (for exact numbers, see Fig. 1). The intervention phase ended as planned after 12 months.

At baseline, GPs had a median age of 52 years (IQR: 44–60), 72% were male and 91.5% worked in a group practice. The patients had a median age of 70 years (IQR: 60–78), 57%

were male. Table 2 gives detailed information on the study population including group comparison. The median incentive paid to the GPs in the intervention group was 637.50 Swiss francs (IQR: 300–1200) (€: 603.-, IQR: 284–1137).

**Primary Outcomes**

After 12 months, the proportions decreased for the process QM HbA1c from 79.0 to 78.3% for the intervention group and from 81.5 to 81.0% for the control group (Fig. 2). The proportions of patients achieving a blood pressure target level below 140/85 mmHg improved from 49.9 to 52.5% for the intervention group, and decreased from 51.2 to 49.0% for the control group (Fig. 2). The odds ratio for the interactive effect of time and intervention over the entire observation period was 1.09 (95% CI: 0.90–1.32,  $p = 0.39$ ) for the process QM of HbA1c and 1.16 (95% CI: 0.99–1.36,  $p = 0.06$ ) for the clinical QM of achieving a blood pressure target level below 140/85 mmHg (Table 3). The detailed results of the logistic regression, and the estimates of the hierarchical random effects for GPs and GP nested in practices, can be found in Additional File 1, Table 4 and Table 5.

**Secondary Outcomes**

After 12 months, the proportions in all four non-incentivized QMs increased slightly for the intervention group, whereas for the control group, only the proportions for cholesterol QMs increased (Fig. 2). The logistic regression analysis revealed that the intervention had a significant spill-over effect on the two process QMs of the secondary outcomes (blood pressure: OR 1.24, 95% CI 1.02–1.50,  $p < 0.05$ , cholesterol: OR 1.17, 95% CI 1.00–1.38,  $p > 0.05$ ) (Table 3).

**Sensitivity Analysis**

For the sensitivity analysis, we included 33 GPs from 16 different practices with 1541 patients with diabetes mellitus. Practice, GP, and patient characteristics were similar to the study population (Additional File 1, Table 6), whereas baseline proportions were lower for non-participants regarding blood pressure QMs (Additional File 1, Fig. 1). The results



Table 2 Baseline Characteristics by Group

	Intervention group	Control group
Practice and GP characteristics		
Practices	21	22
GPs	36	35
GPs per practice	1.0 (1.0–2.0)	1.0 (1.0–2.0)
Network participation (no)	4 (11.1)	2(5.7)
Age	48.0 (42.0–57.3)	54.0 (46.0–63.0)
Gender (male)	24 (66.7)	27 (77.1)
Volume of patients with DM per GP (%)	5.0 (3.3–6.4)	4.9 (3.7–6.4)
Patient characteristics		
Patients	1771	2067
Patient age	70.0 (59.0–79.0)	69.0 (60.0–77.0)
Patient gender (male)	989 (55.8)	1209 (58.6)
Treatment and disease characteristics		
Consultations in observation period	8 (3–15)	7 (3–12)
Blood pressure measurements	3 (1–4)	2 (1–4)
Systolic blood pressure [mm Hg]	137.5 (128.4–149.0)	134.0 (125.0–143.3)
Diastolic blood pressure [mm Hg]	80.0 (74.0–86.0)	80.0 (73.3–85.0)
HbA1c measurements	2 (1–3)	3 (1–4)
HbA1c [%]	6.8 (6.3–7.5)	6.8 (6.3–7.5)
Cholesterol measurement	1 (1–2)	1 (1–1)
Cholesterol [mmol/L]	4.5 (3.8–5.3)	4.6 (3.8–5.4)
BMI measurements	2 (1–3)	2 (1–3)
BMI [kg/m <sup>2</sup> ]	29.4 (26.5–33.1)	29.6 (26.3–33.5)
DM associated comorbidities		
Obesity	759 (42.9)	855 (41.4)
Hypertension	1566 (88.4)	1765 (85.4)
Hyperlipidemia	1098 (62.0)	1167 (56.5)
Chronic kidney disease	212 (12.0)	242 (11.7)
Peripheral arterial disease	54 (3.0)	80 (3.9)
Coronary heart disease	127 (7.2)	107 (5.2)
Heart failure	48 (2.7)	35 (1.7)
Stroke	29 (1.6)	38 (1.8)
Retinopathy	11 (0.6)	3 (0.1)
Neuropathy	44 (2.5)	56 (2.7)
DM associated medication		
Insulin only	159 (9.0)	129 (6.2)
Oral anti-diabetic medication only	866 (48.9)	1112 (53.8)
Combination therapy insulin and oral anti-diabetic medication	258 (14.6)	262 (12.7)
Antihypertensive medication	862 (48.7)	975 (47.2)
Antiplatelet therapy and anticoagulants	736 (41.6)	806 (39.0)
Lipid-lowering medication	258 (14.6)	262 (12.7)

Data are presented as n, n (%), or median (interquartile range), unless otherwise stated. GP general practitioner, DM diabetes mellitus, BMI body mass index. Identification schemes for diseases can be found in Additional File 1, Table 3

of the logistic regressions showed that there was no study independent improvement occurring during the observational period (Additional File 1, Table 7) and that the educational feedback report had no effect.

## DISCUSSION

In this parallel cluster RCT, we evaluated the effect of financial incentives on QMs in the treatment of patients with diabetes mellitus in primary care. We tested financial incentives targeted at the GP level in combination with feedback reports, in comparison to feedback reports only. Financial incentives did not have a significant effect on primary outcomes; the proportion of patients receiving annual HbA1c measurements and the proportion of patients achieving the recommended blood pressure target level was stable. Some effects were observed on secondary outcomes.

Systematic reviews<sup>3,4</sup> concluded that financial incentives targeting process QMs, which can directly be altered by providers, showed greater effects than financial incentives

targeting clinical QMs, which can only be influenced indirectly. In contrast, in this study, we detected no significant effect of the financial incentives on the directly incentivized process and clinical QM. However, in case of the process QM of measuring HbA1c annually, for which rates of 80% were achieved, we suspect a ceiling effect, and therefore, no further improvement was attained. Measuring HbA1c is the most standard procedure in diabetes mellitus care and other quality initiatives already taking place in the Swiss health care environment might have led to such already high-performance levels at baseline<sup>23,24</sup> and notably the fee-for-service setting where the setting itself already gives an incentives for regular measurement. However, other European countries report achievement rates of over 90%,<sup>25</sup> implying that higher rates are not beyond reach. The appearance of a ceiling effect for HbA1c measures is also supported by our models, which reveal smaller variability of random effects on practice and GP level for HbA1c, than for the other QMs. Furthermore, the presence of ceiling effects is supported when comparing to the achievement rates of 2014 where lower achievement rates but higher variation was observed.<sup>26</sup>

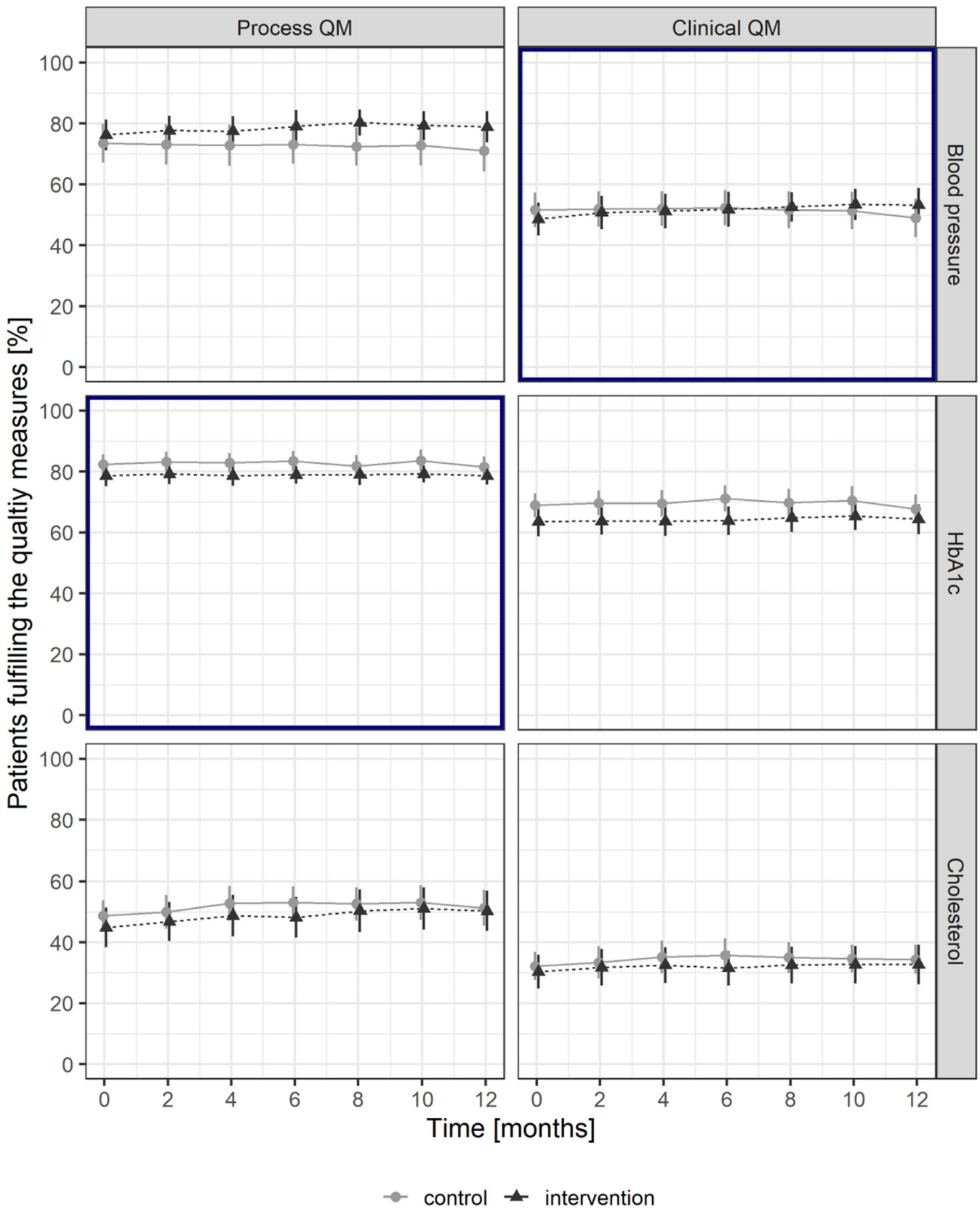


Figure 2 Proportions of patients fulfilling the quality measures during the observation period. Primary outcomes are surrounded by thick frames. Illustrated are mean and Wald 95% confidence interval for each group. QM quality measure.

**Table 3 Interactive Effect of Time and Intervention over the Entire Observation Period**

Type	Subject	OR	95% CI	p value
Primary outcomes				
Clinical QM	Blood pressure	1.16	0.99–1.36	0.06
Process QM	HbA1c	1.09	0.90–1.32	0.39
Secondary outcomes				
Process QM	Blood pressure	1.24	1.02–1.50	<0.05
Clinical QM	HbA1c	1.15	0.97–1.35	0.11
Process QM	Cholesterol	1.17	1.00–1.38	<0.05
Clinical QM	Cholesterol	1.06	0.90–1.25	0.47

The model is adjusted for age and gender of the GPs and the volume of patients with diabetes mellitus per GP. QM quality measure, OR odds ratio, CI confidence interval

Despite we observed no effect in directly incentivized QMs, we observed some effects on non-incentivized process QMs. It is known that spill-over effects of financial incentives on non-incentivized QM can occur, generally with smaller effect sizes than in incentivized QMs.<sup>27</sup> Spill-over effects might indicate that GPs' awareness on the condition improved and that GPs addressed a more holistic disease approach.<sup>28</sup> Hysong et al. also reported that financial incentives improved care documentation without necessarily improving the care provided. Despite these plausible mechanisms of spill-over effects, our finding should be treated with caution, as the effect on directly incentivized QMs—presumably due to ceiling effects—was limited and other unknown confounders might matter.

Our results show that the potential of financial incentives to increase proportions of patients fulfilling a QM might be limited, and that feedback reporting with educational aspects only had no effect. First, the promised incentive of 75 Swiss francs per percentage point improvement was rather low. With an assumed effect size, the amount achieved would have been slightly less than 1% of the annual fee-for-service income of a Swiss GP. Compared to the QOF, where GPs received financial incentives of up to 25% of their income,<sup>29</sup> our incentive was most likely too low to have a major effect. Second, P4P initiatives and feedback reporting are complex and behavioral economic principles and the design aspects play an important role. Factors such as the amount of the incentive, incentive frequencies, and the level of payment have an influence on the effectiveness of incentives.<sup>30–32</sup> Feedback reports are found to be most effective when they are actionable and individualized and when specific goals are set.<sup>33,34</sup>

### Strengths and Limitations of This Study

To our knowledge, this is the first cluster RCT testing a P4P strategy in diabetes care within Europe. In general, quality improvement studies in diabetes care are of importance, due to high prevalence and disease burden. We were able to implement a simple P4P intervention and to blind the control group. Additionally, we were able to compare the study results with another cohort not participating in the study. This study therefore closes a gap, left open by the many observational studies. Compared to the nationwide-conducted and governmentally

authorized intervention in the UK, we were only able to test our intervention with a comparably small number of GPs, however large enough to validly detect potential effects.

The major limitation of this study is its inherent risk of selection bias as GPs participate on a voluntary basis in the FIRE project. However, GPs participating in the FIRE project cover 8% of GPs working in the German-speaking region of Switzerland<sup>35</sup> but they are most likely GPs more engaged in research and quality of care, than their counterparts not participating in the FIRE project.<sup>36</sup> Additionally, using EMR favors higher quality.<sup>37,38</sup> However, EMRs are only used by around 70% of GPs in Switzerland<sup>39</sup> but an EMR is a key requirement for participating in the FIRE project.

In the baseline analysis of this RCT, we conducted further investigations to disentangle the complex interplay of different factors influencing quality of care, also including patient characteristics.<sup>40</sup> However, the influence of available explanatory variables of practice, GP, and patient level on performance was surprisingly small. Another potential explanatory variable—not systematically retrieved—is the availability of specifically trained chronic care nurses. To our knowledge, only one practice had a special trained nurse, since in Switzerland it is highly uncommon that trained nurses work in GP practices.

Further limitations arose due to the underlying database. Such databases are known to be prone to missing data and data quality issues. These issues might not be apparent at baseline due to randomization. However, we cannot preclude that the increases in proportions of patients fulfilling process QMs are due to better data reporting, and not due to enhanced care delivery. This confounder might especially be of concern in the process QM of blood pressure measurement, as it is very likely that the blood pressure values, which need to be entered manually, are not exclusively reported in the intended field. However, cross-validation indicated high validity of the database.<sup>41</sup> Lastly, limited knowledge was available about death of patients, as only very few GPs code it appropriately to be visible for FIRE. To overcome this limitation, patients with no consultation within the last 12 months were excluded due to loss of follow-up.

### CONCLUSION

GP level financial incentives did not result in more frequent HbA1c measurements or in improved blood pressure control. Interestingly, we observed spill-over effects on non-incentivized process QMs. Yet, the mechanism of spill-over effects of financial incentives is largely unclear. In general, however, the potential of financial incentives to improve the quality of diabetes treatment may be limited. In order to potentially maximize the impact of financial incentives, behavioral economic principles should be given greater consideration.

**Acknowledgements:** Our thanks goes to the participating GPs of the FIRE study group of general practitioners.

**Corresponding Author:** Rahel Meier, Dr.; Institute of Primary Care, University of Zurich, Pestalozzistrasse 24, 8091 Zurich, Switzerland (e-mail: rahel.meier@usz.ch).

The online version contains supplementary material available at <https://doi.org/10.1007/s11606-021-06714-8>.

**Author Contribution** RM: methodology, formal analysis, writing - original draft, writing - review and editing

CC: conceptualization, methodology, writing - review and editing, funding acquisition

FV: data curation, methodology, writing - review and editing

LM: methodology, writing - review and editing

OS: conceptualization, writing - review and editing, funding acquisition

TR: conceptualization, writing - review and editing, funding acquisition

**Funding** Open Access funding provided by Universität Zürich. This study was supported by a grant from the Swiss National Science Foundation (Grant Number 407440\_167204).

**Data Availability** The data are gathered within the ongoing FIRE project. The FIRE database can be accessed at any time by the scientific team of the institute. For external requests, access has to be requested from the head of the institute.

#### Declarations:

**Ethics Approval:** According to the local ethics committee of the canton of Zurich, the project does not fall under the scope of the law on human research and therefore, no ethical consent was necessary to conduct the trial (BASEC-Nr. Req-2017-00797).

**Conflict of Interest:** The authors declare that they do not have a conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## REFERENCES

- World Health Organization. Global report on diabetes. Geneva 2016.
- Tricco AC, Ivers NM, Grimshaw JM, Moher D, Turner L, Galipeau J, et al. Effectiveness of Quality Improvement Strategies on the Management of Diabetes: a Systematic Review and Meta-analysis. *Lancet*. 2012;379(9833):2252-61. doi:[https://doi.org/10.1016/S0140-6736\(12\)60480-2](https://doi.org/10.1016/S0140-6736(12)60480-2)
- Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic Review: Effects, Design Choices, and Context of Pay-for-Performance in Health Care. *Bmc Health Serv Res*. 2010;10:247. doi:<https://doi.org/10.1186/1472-6963-10-247>
- Mendelson A, Kondo K, Damberg C, Low A, Motiapaaka M, Freeman M, et al. The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care: a Systematic Review. *Annals of Internal Medicine*. 2017;166(5):341-53. doi:<https://doi.org/10.7326/M16-1881>
- Scott A, Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, et al. The Effect of Financial Incentives on the Quality of Health Care Provided by Primary Care Physicians. *Cochrane Database Syst Rev*. 2011(9):CD008451. doi:<https://doi.org/10.1002/14651858.CD008451.pub2>
- Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of Pay for Performance in Health Care: a Systematic Review of Systematic Reviews. *Health Policy*. 2013;110(2):115-30. doi: <https://doi.org/10.1016/j.healthpol.2013.01.008>
- Kontopantelis E, Springate DA, Ashworth M, Webb RT, Buchan IE, Doran T. Investigating the Relationship Between Quality of Primary Care and Premature Mortality in England: a Spatial Whole-Population Study. *BMJ*. 2015;350:h904. doi:<https://doi.org/10.1136/bmj.h904>
- Forbes LJ, Marchand C, Doran T, Peckham S. The Role of the Quality and Outcomes Framework in the Care of Long-Term Conditions: a Systematic Review. *Br J Gen Pract*. 2017;67(664):e775-e84. doi:<https://doi.org/10.3399/bjgp17X693077>
- Chang RE, Lin SP, Aron DC. A Pay-for-Performance Program in Taiwan Improved Care for Some Diabetes Patients, but Doctors May Have Excluded Sicker Ones. *Health Aff (Millwood)*. 2012;31(1):93-102. doi:<https://doi.org/10.1377/hlthaff.2010.0402>
- Chen TT, Chung KP, Lin IC, Lai MS. The Unintended Consequence of Diabetes Mellitus Pay-for-Performance (P4P) Program in Taiwan: Are Patients with More Comorbidities or More Severe Conditions Likely to be Excluded from the P4P Program? *Health Serv Res*. 2011;46(1 Pt 1):47-60. doi:<https://doi.org/10.1111/j.1475-6773.2010.01182.x>
- Petersen LA, Simpson K, Pietz K, Urech TH, Hysong SJ, Profit J, et al. Effects of Individual Physician-Level and Practice-Level Financial Incentives on Hypertension Care: a Randomized Trial. *JAMA*. 2013;310(10):1042-50. doi:<https://doi.org/10.1001/jama.2013.276303>
- Bardach NS, Wang JJ, De Leon SF, Shih SC, Boscardin WJ, Goldman LE, et al. Effect of Pay-for-Performance Incentives on Quality of Care in Small Practices with Electronic Health Records: a Randomized Trial. *JAMA*. 2013;310(10):1051-9. doi:<https://doi.org/10.1001/jama.2013.277353>
- Asch DA, Troxel AB, Stewart WF, Sequist TD, Jones JB, Hirsch AG, et al. Effect of Financial Incentives to Physicians, Patients, or Both on Lipid Levels: a Randomized Clinical Trial. *JAMA*. 2015;314(18):1926-35. doi:<https://doi.org/10.1001/jama.2015.14850>
- Navathe AS, Volpp KG, Caldarella KL, Bond A, Troxel AB, Zhu J, et al. Effect of Financial Bonus Size, Loss Aversion, and Increased Social Pressure on Physician Pay-for-Performance: a Randomized Clinical Trial and Cohort Study. *JAMA network open*. 2019;2(2):e187950. doi:<https://doi.org/10.1001/jamanetworkopen.2018.7950>
- OECD/WHO. OECD Reviews of Health Systems: Switzerland 2011.
- Chmiel C, Bhend H, Senn O, Zoller M, Rosemann T, study-group F. The FIRE Project: a Milestone for Research in Primary Care in Switzerland. *Swiss Med Wkly*. 2011;140:w13142. doi:<https://doi.org/10.4414/smw.2011.13142>
- WONCA ICC. International Classification of Primary Care - ICPC-21998.
- Meier R, Muheim L, Senn O, Rosemann T, Chmiel C. The Impact of Financial Incentives to Improve Quality Indicators in Patients with Diabetes in Swiss Primary Care: a Protocol for a Cluster Randomised Controlled Trial. *BMJ Open*. 2018;8(6):e023788. doi:<https://doi.org/10.1136/bmjopen-2018-023788>
- WHO Collaborating Centre for Drug Statistics Methodology. ATC-Structure and principles. Oslo, Norway 2018. [https://www.whocc.no/atc/structure\\_and\\_principles/](https://www.whocc.no/atc/structure_and_principles/). Accessed 09.07.2019.
- Künzi K SS, Stocker D. Erhebung der Einkommensverhältnisse der berufstätigen Ärzteschaft. *Schweizerische Ärztezeitung*. 2011(36).
- Teerenstra S, Moerbeek M, van Achterberg T, Pelzer BJ, Borm GF. Sample Size Calculations for 3-Level Cluster Randomized Trials. *Clin Trials*. 2008;5(5):486-95. doi:<https://doi.org/10.1177/1740774508096476>
- Grischott T. The Shiny Balancer - Software and Imbalance Criteria for Optimally Balanced Treatment Allocation in small RCTs and cRCTs. *BMC medical research methodology*. 2018;18(1):108. doi:<https://doi.org/10.1186/s12874-018-0551-5>
- Huber CA, Brandl M, Rapold R, Reich O, Rosemann T. A Set of Four Simple Performance Measures Reflecting Adherence to Guidelines Predicts Hospitalization: a Claims-Based Cohort Study of Patients with Diabetes. *Patient Prefer Adherence*. 2016;10:223-31. doi:<https://doi.org/10.2147/PPA.S99895>



24. Stefan Heini, Glavic D. *Helsana-Standpunkt, Magazin zur Gesundheits- und Unternehmenspolitik der Helsana-Gruppe*. Zürich: Helsana Gruppe 2017.
25. **Stone MA, Charpentier G, Doggen K, Kuss O, Lindblad U, Kellner C, et al.** Quality of Care of People with Type 2 Diabetes in Eight European Countries: Findings from the Guideline Adherence to Enhance Care (GUIDANCE) Study. *Diabetes care*. 2013;36(9):2628-38. doi:<https://doi.org/10.2337/dc12-1759>
26. **Djalali S, Frei A, Tandjung R, Baltensperger A, Rosemann T.** Swiss Quality and Outcomes Framework: Quality Indicators for Diabetes Management in Swiss Primary Care Based on Electronic Medical Records. *Gerontology*. 2014;60(3):263-73. doi:<https://doi.org/10.1159/000357370>
27. **Doran T, Kontopantelis E, Valderas JM, Campbell S, Roland M, Salisbury C, et al.** Effect of Financial Incentives on Incentivised and Non-incentivised Clinical Activities: Longitudinal Analysis of Data from the UK Quality and Outcomes Framework. *BMJ*. 2011;342:d3590. doi:<https://doi.org/10.1136/bmj.d3590>
28. **Hysong SJ, SoRelle R, Broussard Smitham K, Petersen LA.** Reports of unintended consequences of financial incentives to improve management of hypertension. *Plos One*. 2017;12(9):e0184856. doi:<https://doi.org/10.1371/journal.pone.0184856>
29. Roland M, Guthrie B. Quality and Outcomes Framework: What Have we Learnt? *BMJ (Clinical research ed.)*. 2016;354:i4060-i. doi:<https://doi.org/10.1136/bmj.i4060>
30. **Emanuel EJ, Ubel PA, Kessler JB, Meyer G, Muller RW, Navathe AS, et al.** Using Behavioral Economics to Design Physician Incentives That Deliver High-Value Care Behavioral Economics, Physician Incentives, and High-Value Care. *Annals of Internal Medicine*. 2016;164(2):114-9. doi:<https://doi.org/10.7326/M15-1330>
31. **Doran T, Maurer KA, Ryan AM.** Impact of Provider Incentives on Quality and Value of Health Care. *Annual Reviews*. 2017;38(1):449-65. doi:<https://doi.org/10.1146/annurev-publhealth-032315-021457>
32. **Bond AM, Volpp KG, Emanuel EJ, Caldarella K, Hodlofski A, Sacks L, et al.** Real-time Feedback in Pay-for-Performance: Does More Information Lead to Improvement? *J Gen Intern Med*. 2019;34(9):1737-43. doi:<https://doi.org/10.1007/s11606-019-05004-8>
33. **Hysong SJ.** Meta-analysis: Audit and Feedback Features Impact Effectiveness on Care Quality. *Medical care*. 2009;47(3):356-63. doi:<https://doi.org/10.1097/MLR.0b013e3181893f6b>
34. **Hysong SJ, Best RG, Pugh JA.** Audit and Feedback and Clinical Practice Guideline Adherence: Making Feedback Actionable. *Implementation Science*. 2006;1(1):9. doi:<https://doi.org/10.1186/1748-5908-1-9>
35. Association FSM. Online query tool of the Swiss physician statistics [Online-Abfrage tool der FMH-Ärztstatistik]. 2020. <http://aerztestatistik.myfmh2.fmh.ch>. Accessed 26.02.2020 2020.
36. **Dahrouge S, Armstrong CD, Hogg W, Singh J, Liddy C.** High-Performing Physicians are More Likely to Participate in a Research Study: Findings from a Quality Improvement Study. *BMC Medical Research Methodology*. 2019;19(1):171. doi:<https://doi.org/10.1186/s12874-019-0809-6>
37. **Riordan F, McHugh SM, O'Donovan C, Mtshede MN, Kearney PM.** The Role of Physician and Practice Characteristics in the Quality of Diabetes Management in Primary Care: Systematic Review and Meta-analysis. *J Gen Intern Med*. 2020;35(6):1836-48. doi:<https://doi.org/10.1007/s11606-020-05676-7>
38. Kern LM, Barrón Y, Dhopeswarkar RV, Edwards A, Kaushal R. Electronic Health Records and Ambulatory Quality of Care. *J Gen Intern Med*. 2013;28(4):496-503. doi:<https://doi.org/10.1007/s11606-012-2237-8>
39. **Pahud O.** eHealth in der ambulanten Grundversorgung. *Obsan Bulletin* 3/2020 - Swiss Health Observatory. 2020.
40. **Meier R, Valeri F, Senn O, Rosemann T, Chmiel C.** Quality Performance and Associated Factors in Swiss Diabetes Care – a Cross-sectional Study. *Plos One*. 2020;15(5):e0232686. doi:<https://doi.org/10.1371/journal.pone.0232686>
41. **Zellweger U, Bopp M, Holzer BM, Djalali S, Kaplan V.** Prevalence of Chronic Medical Conditions in Switzerland: Exploring Estimates Validity by Comparing Complementary Data Sources. *BMC Public Health*. 2014;14(1):1157. doi:<https://doi.org/10.1186/1471-2458-14-1157>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.