ORIGINAL ARTICLE

Open Access

On the effectiveness of case management for people with disabilities

Matthias Draheim[1], Peter Schanbacher[2] and Ruben Seiberlich[3]*

Abstract

Case managers provide individual and comprehensive support to employees who have become incapable of working. Using data from a large insurance company we find that overall, 43.9% of the people in our sample could be reintegrated. Controlling for personal characteristics, we analyze the effectiveness of case management by modelling the probability of reintegrating people being incapable of working into the labor market. Using parametric and semiparametric decomposition methods, we control for observational differences. We analyze how much of the difference in the reintegration rate between people who participate in case management and those who do not, is due to differences in characteristics and how much is due to case management itself. We find that the estimated probability of reintegration is 18.9% higher if people participate in case management. Moreover, our results show that no more than 15% are due to differences in characteristics and at least 85% can be attributed to case management itself.

Keywords: Return to work, Insurance, Case management, Labor force participation, Individual support on reintegration, Logistic regression, Decomposition, Propensity score matching

JEL classification: C14, G22, I11, I13

1 Introduction

Labor participation is a precondition to the social and economic stability of any developed country. Health status is the main driver of the ability of workers to be employed (Ferlie et al. 2016). Often, poor health conditions lead to a disability pension (Robroek et al. 2013), unemployment (Cardano et al. 2004 and Schuring et al. 2007) or early retirement (van Berg et al. 2010). In the current economic condition in western societies with low unemployment and an aging society, governments have strong incentives to implement policies to reintegrate workers with a disability as soon as possible. For the US the economic estimated lifetime savings for removing an additional one percent of the beneficiaries from the rolls of the disability insurance and the supplemental security income programs each year are several billion USD (U.S. General Accounting Office 1997). In the US alone, there are over eight million disability insurance beneficiaries with average yearly payments of more than USD 6000 (Social Security Administration 2013) with rising expenditures for the government (Segelken 2014). Several return-to-work interventions have been started by the OECD countries (see Clayton et al. 2012 for a systematic review as well as Burkhauser et al. 2016, for trends in government disability programs).

For the person with a disability, employment itself is positive in several dimensions. It positively impacts the general social status and professional skills. The sense of engagement, social identification, social interactions with other employees as well as mental stimulation support the general well-being (Noh et al. 2015 or Waddell and Bruton 2006). In case of (temporary) disability reintegration improves the health (Schuring et al. 2011).

Individual support of professionals, specialized in rehabilitation methods with a supportive attitude, does improve the chance of reintegration into the workforce (see Liukko and Kuuva 2016; or Hansen 2006 for a study

*Correspondence: ruben.seiberlich@zhaw.ch
[3] ZHAW School of Management and Law, St.-Georgen-Platz 2, 8401 Winterthur, Switzerland
Full list of author information is available at the end of the article

using German data). Improvements to the work environment by the employer is an important factor for reintegration (Wheeler et al. 2001).

Whether a person has been successfully reintegrated into the labor market can be viewed differently by the different stakeholders (Young et al. 2005). For the insurance company the reintegration has been successful if the benefit payments can be stopped. An employer sees successful reintegration when the person is fully able to work again. The employee feels successfully reintegrated if she can continue her employment permanently and for the same salary. In our analysis a person has a disability when she cannot work at least 40% for 1 year. Accordingly, the reintegration is successful if the person is more than 60% fit for work again (measured by the comparison of income with and without disability). Claims for both physical and psychological reasons are covered. The definition corresponds to Swiss practice and legal principles (see Swiss Federal Law on Disability Insurance, Article 28).

To analyze the effectiveness of case management we use exclusive data. The data is unique in several respects. It is based on more than 42,000 cases of incapacity for work in the Swiss labor market and provides information on personal characteristics such as gender, age, salaries, date, and cause of the disability. It is known that personal characteristics are more relevant to reintegration than employer characteristics (see Galizzi et al. 2019). The Swiss labor market, in contrast to most other economies is particularly interesting for such an analysis. In Switzerland, employers or associated pension funds to be exact must cover incapacity to work, as well as death and old age, at a high level. This makes it possible to consider not only a selective group of privileged employees in large corporations, but the entire working population. Due to the high obligatory coverage, the disability benefits within all workers are comparable at about 60% of the initial wage. There are no large differences in disability benefits between workers (relative to the income). Opposed to other countries, in Switzerland there are no low-paid workers losing most of their income in case of disability. The risk of incapacity to work is borne by independent pension funds or insurance companies. We obtain the data from the largest semi-autonomous insurer in Switzerland.[1] Government reintegration measures are often moderately recorded and less financially monitored. Due to financial incentives the statistics of private insurance companies are more reliable. The insurance companies have an interest in knowing and managing their portfolio

comprehensively to determine annual premiums. If an insured person becomes unable to work, the insurer may offer the person a so-called case manager for support. The offer from the insurer as well as the acceptance from the insured is on a voluntary basis. Participation in case management has no effect on benefit payments and no additional costs for the insured person. The case manager provides emotional, legal, and medical support to give the insured person the best possible chances for reintegration into the labor market.

Based on the unique dataset, we analyze the effectiveness of case management, on the probability of reintegration into the labor market. Our results show that if employees, who are not able to work, receive support at an early stage of their disability, the chance of reintegration into the labor market can be significantly increased by about 18.9%.

The rest of the paper is structured as follows. After a detailed explanation of case management, we discuss the data and present the methods used. The results and robustness checks are given in section five. Section six summarizes our findings and concludes.

2 Case management

In an event of incapacity to work the insured person, the employer as well as the insurance company look for a quick recovery and fast reintegration into the labor market. Comprehensive support for the person with a disability belongs to the core competences of the insurer. This takes place in the form of so-called case management (sometimes also called care management). The insurance company decides whether case management is offered. Typically, case management is offered to people with a disability that is not likely to be self-resolving, often these are, for example, disabilities caused by pregnancy complications or minor accidents. The insurance company generally focuses on incapacity to work due to a mental cause combined with a high benefit amount. In such cases, a case manager is most likely to save the insurance company a large loss. The case manager provides individual and comprehensive support to a person with a disability. For the persons concerned the participation in case management is voluntary and in any case without loss of benefits. They can participate or reject the program without further costs and without the loss of benefits. Case management typically takes place in a trusting cooperation of the case manager and the person with a disability.

The support is comprehensive in professional, medical, social, and legal terms. Medical support is provided in the form of coordination of medical measures with internal and external medical specialists to obtain a view of the potential treatment alternatives. On the occupational level case managers help to find a solution with

---

[1] Semi-autonomous insurers cover the risk of death and disability only but not the risk of longevity.

the employer to avoid a replacement and to optimize the work situation for the insured person. Thereby, the case manager focuses on evaluating alternatives and helps to organize further training or retraining. The case manager also analyzes the current social network with the primary goal to build up a supportive environment, to provide emotional support and to break down old, stressful structures. Legally, the case manager coordinates the services between the various insurers (state, private, professional) and assists the affected person with advice, relying on the support of the legal services of the insurance company.

In the first meeting of the case management process, one typically defines the agenda of the case management program and immediate measures are initiated. In a second step, the situation is evaluated through self-assessment and external assessment. New measures are introduced based on these findings. In the third step, a goal is formulated together with the person with a disability. Actions are planned accordingly, and professional offices are involved (doctors, psychologists, lawyers, etc.). Finally, implementation is closely monitored and regularly reviewed. The frequency depends on the cause of disability. In the case of an accident the case management period is shorter, however more frequent (e.g. monthly check of the physiotherapy results). In case of cancer, the case management process lasts longer but meetings of the person with a disability and the case manager are less frequent (e.g. an annual consultation of the preventive examinations). After an overall assessment has been made, case management is concluded.

3 Data

We consider data of a large Swiss life insurer. In Switzerland, pension funds have to cover longevity, mortality, and disability risks in the context of the occupational pension provision system. Many of these pension funds have to insure against these risks according to Swiss regulation or wish to do so due to their small size and the associated lack of diversification. The corresponding portfolio of the insurer is called group life portfolio. The database covers data on 42,361 insured persons (also called cases) having been exposed to a disability between January 1, 2009 and September 30, 2018.

Following Wacholder et al. (1992a) we define the base as the members of the source population during the time periods when they are eligible to become cases. The Swiss occupational pension scheme allows the insured company to choose a waiting period for premium waiver for their employees who become disabled. Most of the time this is three to six months. Since the insured company can report an incapacity of an employee already before the end of the waiting period, several people in the data are only eligible for case

management after 6 months. Wacholder et al. (1992b) highlight that the goal in control selection is to obtain a sample of eligible subjects. As we cannot observe the waiting period, we apply the filter of 180 days to ensure that the control selection is a sample of eligible subjects only. This allows us to estimate the effectiveness of case management for people who are generally eligible for case management. Additionally, we also exclude 494 cases because of missing information. This leaves us with data on 18,237 insured people.

All these insured persons are beneficiaries of at least one of the pension funds being reinsured by the Swiss life insurer. We observe the gender ("Male"), age in years ("Age"), the mean salary in CHF 1000 before the occurrence of the disability ("Income"), the time in years between the beginning of the disability and the date when information on disability is reported to the insurer ("Years since notification"), a flag if the disability is due to cancer ("Cancer"), and the cause of the disability—for each cause there is a dummy which is 1 if the disability is due to this cause and 0 otherwise. "Cancer" is not considered as a cause of the disability itself but always has an interaction with a cause code. For instance, for an insured person suffering from testicular cancer, "Cause: Genitals" is 1 and "Cancer" is 1. During chemotherapy most cancer patients cannot work (de Boer et al. 2008). But even for milder symptoms, such as fatigue, the disease often leads to a strong impairment of the ability to work (Hofman et al. 2007). A low reintegration rate can be expected. Whenever there is no cause code for an insured person available, "Cause: Others" is set to 1 for that person. Moreover, we control for the time in years between 01/01/2009 and the date when disability occurs ("Time in years since 01/01/2009"). This variable controls for time effects, which is needed to account for regulatory changes and differences in interest rate levels. Differences in interest rate levels impact, for example, the costs of the case through changes in the discounted disability payments.

The case management unit of the Swiss insurer provides us with data on the case management cases the unit has dealt with. Merging the case management data with the data on 18,237 insured persons, we find that 404 of 18,237 insured persons have participated in case management. Out of these 404 insured people, participating in case management, 210 or 52.0% could reintegrate to the labor market. Compared to this only 43.7% of those who had not participated in case management could reintegrate to the labor market.

Our descriptive statistics show that overall, 43.9% of the people concerned could be reintegrated. Figure 1 presents the reintegration rates for each of the different causes for disability. A chi-squared test rejects the null
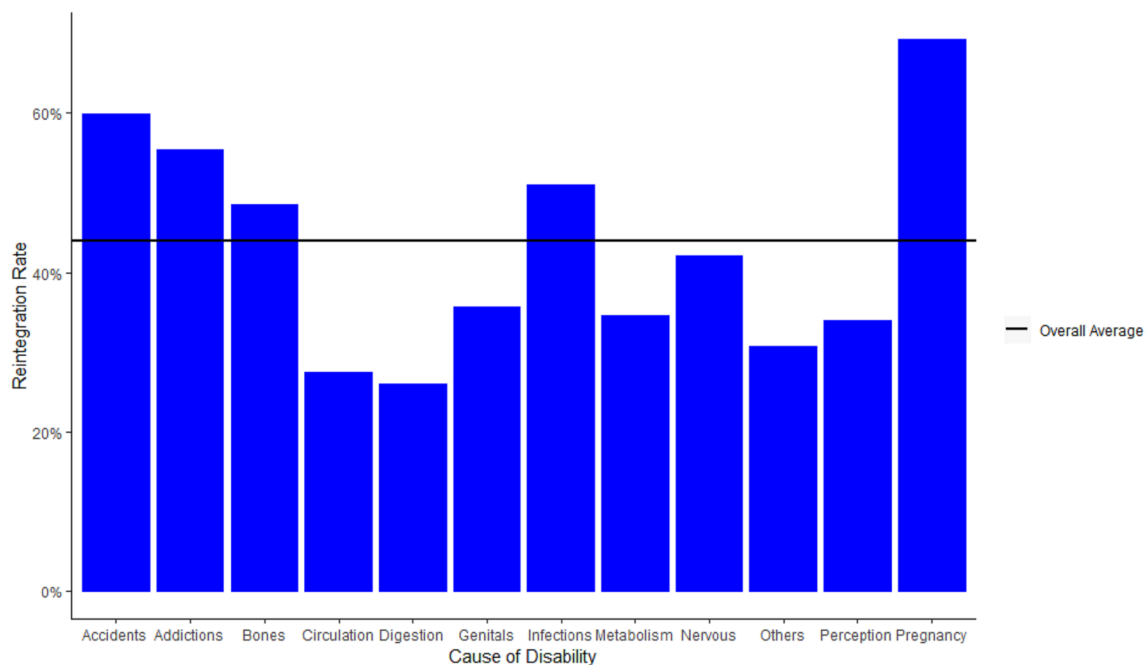
Fig. 1 Reintegration rates per cause of disability

hypothesis that the reintegration rates are independent of the causes for disability with a p-value < 0.001.

The highest reintegration rate is found for people with a disability due to complications in pregnancy (69.2%), followed by those who have an accident (59.9%) and people with an addiction (55.4%). The lowest reintegration rates are found for people with a disability due to digestion (26.0%) or circulation (27.5%).

The variable of interest in this study is a flag indicating if an insured person got reintegrated. An insured person is said to be reintegrated if she can resume working more than 60% of the level before the disability. For calculation the income is measured before and after the disability. Hence, she might be reintegrated by resuming work at another workplace. Table 1 displays the minimum, maximum, the total average, the average of insured who have been reintegrated and those who have not as well as the t-statistic of the respective difference-in-means test.

The table shows that most variables have significantly different means for reintegrated vs. not reintegrated persons. The first row shows that 2.6% of the people who have been reintegrated, have been part of a case management program, whereas only 1.9% of the people who have not been reintegrated have participated in a case management program. The last column shows that this difference is highly statistically significant. This points to a positive impact of case management on reintegration rates. The share of women is significantly higher for the reintegrated group.

Reintegrated persons are on average around 45 years old, while the average age of the remainder is around 49 years. The difference is strongly significant. Apparently, younger people can be reintegrated more easily. The difference in salaries is not significantly different. Hence, the rate of integration seems to be uncorrelated with the level of salaries. The date of the disability plays a significant role, the date of the disability is on average significantly earlier for those who have been reintegrated. On the one hand, a person being exposed to a disability earlier than another person has more time to reintegrate, on the other hand, that might reflect regulatory frameworks or economic conditions changing over time. As expected, it has a significant impact when the insurer gets information on disabilities such that it can start the case management process, the notification period is significantly shorter for the reintegrated group. Among the reintegrated persons, the share of those having cancer is significantly smaller in comparison to the persons who could not be reintegrated which is in line with our expectations due to the severity of this diagnosis. There are significant differences for all causes of the disabilities, except for infections.

4 Empirical strategy

To estimate the effectiveness of case management on reintegration into the labor market we first estimate a multivariate logistic regression controlling for several personal characteristics. Let Y_i be a dummy variable, which is 1 if individual i has reintegrated and 0 otherwise.

Table 1 Descriptive statistics and tests

Variable	Min	Max	All	Reint = 1	Reint = 0	Test st
Case Management	0	1	2.2%	2.6%	1.9%	−3.2
Male	0	1	62.4%	61.4%	63.1%	2.3
Age (in years)	17.24	64.52	47.29	44.76	49.27	27.3
Income (in CHF 1000)	1	774	68.56	68.12	68.91	1.4
Time in years since 01/01/2009	0	9.48	4.84	4.39	5.20	22.3
Years since notification	0	8.79	0.67	0.52	0.79	27.1
Cancer	0	1	10.3%	6.2%	13.5%	16.9
Cause: Circulation	0	1	7.0%	4.4%	9.1%	12.8
Cause: Digestion	0	1	4.4%	2.6%	5.8%	11.0
Cause: Genitals	0	1	4.8%	3.9%	5.5%	5.1
Cause: Nervous	0	1	31.4%	30.2%	32.4%	3.3
Cause: Perception	0	1	1.8%	1.4%	2.1%	3.8
Cause: Infections	0	1	0.3%	0.3%	0.2%	−1.0
Cause: Bones	0	1	23.5%	26.0%	21.6%	−6.9
Cause: Metabolism	0	1	1.8%	1.4%	2.1%	3.6
Cause: Accidents	0	1	17.0%	23.2%	12.2%	−19.3
Cause: Addictions	0	1	0.4%	0.5%	0.3%	−1.9
Cause: Pregnancy	0	1	0.9%	1.5%	0.5%	−6.3
Cause: Others	0	1	6.5%	4.5%	8.0%	9.8
Number of observations			18,237	8,010	10,227	

The variable "Years since notification" measures the time in years between the beginning of the disability and the date when the information on disability was reported to the insurer

X_i is a vector containing the individual characteristics of individual i including a constant term, and CM_i a dummy variable which is 1 if individual i participates in case management and 0 otherwise. In this case the conditional probability of reintegration into the labor market is given by

$$P(Y_i|CM_i, X_i) = F\left(CM_i\gamma + X_i'\beta\right) \tag{1}$$

where $F(\cdot)$ is the cumulative distribution function of the logistic distribution.

To allow for individual heterogeneity and to relax the functional form assumption in (1) we, additionally, apply propensity score matching to estimate the causal effect of case management on the probability of reintegration into the labor market. Propensity score matching controls for observational differences by comparing the reintegration rates of people with identical characteristics. It has been developed to estimate treatment effects (Rubin 1974; Rosenbaum and Rubin 1983 and 1985; Heckman et al. 1997 and 1998; Dehejia and Wahba 1999; Abadie and Imbens 2006).

Using propensity score matching, we estimate the average treatment effect on the treated (ATT), defined as

$$E[Y_i^1 - Y_i^0|CM_i = 1]$$

where Y_i^1 is the potential reintegration rate for individual i if she participates in case management. Y_i^0 is the potential reintegration rate of individual i if she does not participate in case management and CM_i the treatment variable, which is one if individual i participates in case management and zero otherwise.

The identification based on propensity score matching crucially relies on three assumptions. The first one is the conditional independence assumption (CIA) which guarantees that conditional on confounding variables the potential outcomes are stochastically independent of the treatment: $Y_i^0 \perp CM_i|P_i(X)$, where $P_i(X)$ denotes the propensity score based on the confounding variables of individual i. The CIA requires that all confounding factors associated with the potential outcomes as well as the treatment status are observed. To justify the CIA, we control for a range of covariates available in the data set, including gender, age, income, time effects, time since notification, cancer and cause of disability. The more informative the data are, the easier it is to justify the CIA assumption. Given the data available to us we control for the most important

cofounding variables. On the other hand, we do not observe the motivation of the people nor their employment / unemployment history, exact job description or extent of disability regarding the occupational tasks. Even though important missing variables do correlate with the income, which we control for, the CIA is a strong assumption in our setting. Moreover, Fortin et al. (2011) point out that the aggregate decomposition would even be valid in the presence of the correlation between unobserved and observed characteristics under the condition that the correlation is the same for both groups.

The second assumption, referred to as the Stable Unit Treatment Value Assumption (SUTVA), rules out any influence of an individual's treatment status on another individual's potential outcome (see Rubin 1986). This assumption requires that the potential outcome observation on one person should be unaffected by the assignment of case management to the other units. If SUTVA is violated causal inference becomes more difficult. In our setting the case management is done for everyone on a one-by-one basis and, therefore, the potential outcome observation on one person should be unaffected by the assignment of case management to other people. One crucial assumption, however, is that there is no capacity restriction for the case managers. In case of capacity limitations, the case managers would need to do a selection for case management and therefore the outcome of one person would be affected by the assignment of case management to another person. Going forward we assume that SUTVA holds.

The third assumption is the overlap assumption. It requires that the probability of obtaining case management is smaller than one, i.e. $P(D = 1|X) < 1$. This type of overlap assumption is standard in the literature (e.g. Rosenbaum and Rubin 1983; Heckman et al. 1997; Hahn 1998; Wooldridge 2002; Imbens 2004). There is a stronger version of the overlap assumption called strict overlap (e.g. Robins et al. 1994; Abadie and Imbens 2006; Crump et al. 2009). Strict overlap requires that the probability of having case management is strictly smaller than $1 - \xi$ for some $\xi > 0$. Khan and Tamer (2010) point out that something like the strict overlap assumption is needed for $\sqrt{N}$—convergence of some semiparametric estimators. Busso et al. (2008) provide further evidence on the importance of the (strict) overlap assumption. To guarantee that $supp(X|D = 1) \subseteq supp(X|D = 0)$ we restrict the estimation of the composition and return effect to the common support subpopulation.

Thus, we take advantage of the huge sample size of people without case management by using their characteristics and reintegration rates to determine what the reintegration rate of those with case management would have been if they had not participated in case management. Using propensity score matching we determine, for each person with case management, the people who have most similar characteristics and weight their reintegration rate in the calculation of the counterfactual outcome highest. This allows us to estimate the reintegration rate that people who participate in case management ($CM = 1$) would have, if they did not participate in case management ($CM = 0$), i.e. $E[Y^0|CM = 1]$.

To obtain a causal estimate of the impact of case management, we compare this counterfactual outcome to the expected reintegration rate of the observations with case management, e.g. $E[Y^1|CM = 1]$.

In order to compare only individuals that are comparable, we restrict the sample to the common support subpopulation S. We define the common support as those individuals who have a propensity score lower than the maximum propensity score in the group of people without case management and higher than the propensity score in the group of people with case management, i.e. $S = \left\{ \widehat{p}_i \in \left[\widehat{p}^{min_{CM=1}}; \widehat{p}^{max_{CM=0}} \right] \right\}$. Lechner and Strittmatter (2019) show that this procedure consistently outperforms other procedures to determine the common support. Figure 2 shows the densities of the propensity scores for the two groups. In general, the propensity scores are low for both groups. Those who participate in case management also tend to have a higher propensity score to participate in case management. Moreover, the figure shows that there are controls for those who participate and that a common support is given.

Thereby, $E_S[Y^1|CM = 1]$ is equal to the sample mean of the reintegration rate for those in the common support with case management, denoted by $\overline{Y}_{S,CM=1}$. We apply nonparametric Kernel matching using propensity scores to estimate the expected counterfactual outcome $E_S[Y^0|CM = 1]$. To obtain the propensity scores, we estimate the probability that an individual attends in case management (CM = 1) by a logistic regression, that is $\widehat{p} = \widehat{P}[CM = 1|X = x] = F\left(x'\widehat{\beta}\right)$, where $F(\bullet)$ describes the cumulative logistic distribution.

Let $f_1^S(p)$ be the distribution of the propensity scores $P = P(X)$ for people with case management ($CM = 1$) on the common support. Frölich (2007) shows that the counterfactual outcome is identified as follows:

$$E_S\left[Y^0|CM = 1\right] = \int E[Y|P(X) = p, CM = 0]f_1^S(p)dp.$$
(2)

We estimate the expected outcome at each point $P(X) = p$ using the nonparametric Nadaraya-Watson estimator (Nadaraya, 1964; Watson, 1964)

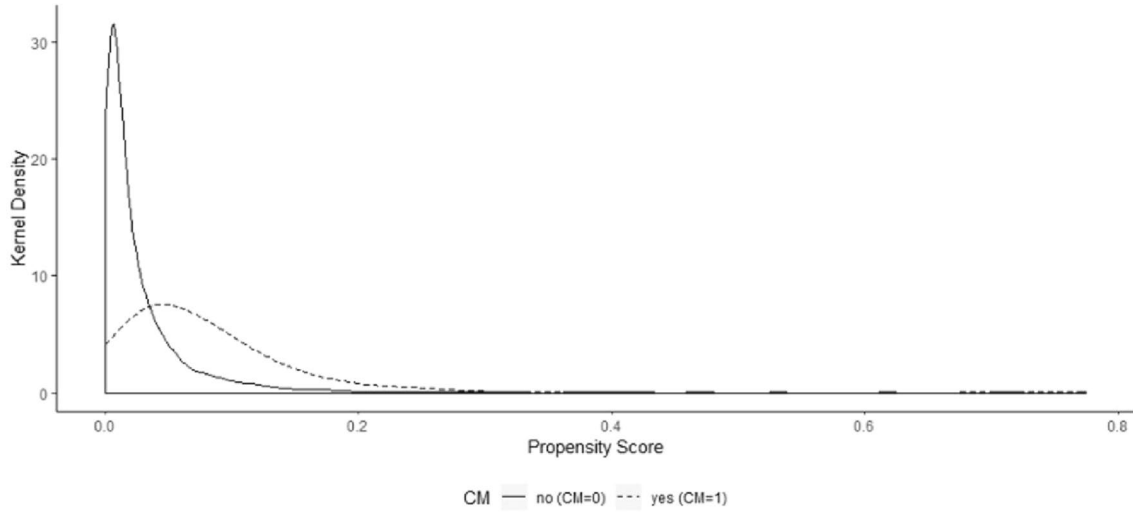**Fig. 2** Density of propensity scores for common support subpopulation

$$\widehat{E}[Y|P(X)=p, CM=0] = \frac{\sum_{i=1}^{n_0} K\left(\frac{p_i^0 - p}{h}\right) * Y_i}{\sum_{i=1}^{n_0} K\left(\frac{p_i^0 - p}{h}\right)} \quad (3)$$

K describes the kernel function, h is the bandwidth, n_0 is the number of observations and p_i^0 the propensity scores of those individuals who have no case management ($CM = 0$). We use a Gaussian Kernel, i.e. $K(\bullet) = \Phi(\bullet)$ and the bandwidth is chosen using Silverman's rule of thumb, i.e. $h = 1.06\widehat{\sigma}n_1^{-1/5}$, (Silverman 1986).

To understand the difference in the reintegration rate between those who participate in case management and those who do not participate, we apply the Blinder-Oaxaca decomposition (Blinder 1973; Oaxaca 1973) for nonlinear models. Following Bauer and Sinning (2008), we decompose the observed difference in the reintegration rates between those who participate in case management and those who do not participate into a part that is due to differences in characteristics between the two groups and into a part that is due to differences in coefficients

characteristics including a constant term, but excluding the dummy variable for case management. The number of observations in the group with case management is denoted by n_1 and n_0 is the number of observations in the group without case management. $F(\bullet)$ denotes the cumulative logistic distribution. Therefore, $\frac{1}{n_1}\sum_{i=1}^{n_1} F\left(X_{i,CM=1}' \widehat{\beta}_{CM=1}\right)$ is simply the sample average of the participation rate for those participating in case management and equal to $\frac{1}{n_1}\sum_{i=1}^{n_1} Y_{i,CM=1}$. Likewise, $\frac{1}{n_0}\sum_{i=1}^{n_0} F\left(X_{i,CM=0}' \widehat{\beta}_{CM=0}\right)$ is equal to $\frac{1}{n_0}\sum_{i=1}^{n_0} Y_{i,CM=0}$ and the sample average of the participation rate for those not participating in case management.

The first term in (4), captures the difference that is due to differences in the coefficients. Our primary interest is in this term, as it shows how the estimated average in the reintegration rate would change, if people with their predefined characteristics had case management. The second term is the difference that is due to differences in characteristics.

To avoid the functional form assumption and to allow for individual heterogeneity, we again rely on propen-

$$\widehat{\Delta} = \underbrace{\frac{1}{n_1}\sum_{i=1}^{n_1} F\left(X_{i,CM=1}' \widehat{\beta}_{CM=1}\right) - \frac{1}{n_1}\sum_{i=1}^{n_1} F\left(X_{i,CM=1}' \widehat{\beta}_{CM=0}\right)}_{\Delta \text{ in coefficients (impact case management)}} + \underbrace{\frac{1}{n_1}\sum_{i=1}^{n_1} F\left(X_{i,CM=1}' \widehat{\beta}_{CM=0}\right) - \frac{1}{n_0}\sum_{i=1}^{n_0} F\left(X_{i,CM=0}' \widehat{\beta}_{CM=0}\right)}_{\Delta \text{ in characteristics}} \quad (4)$$

where $\widehat{\beta}_{CM=k}$ are the estimated coefficients obtained from separate logistic regressions for the group that does participate in case management ($k = 1$) and the group that does not ($k = 0$). $X_{i,CM=k}$ are the corresponding

sity score matching. Following Frölich (2007) we use the counterfactual outcome from (2),

$$E_S\left[Y^0|CM=1\right] = \int E[Y|P(X)=p, CM=0]f_1^S(p)dp,$$

to derive a semiparametric alternative to the Blinder-Oaxaca decomposition for logistic regressions. Therefore, we re-write the difference in the mean reintegration rates between those who attend case management and those who do not as

$$\Delta_S = E_S\left[Y^1|CM = 1\right] - E_S\left[Y^0|CM = 0\right]$$

$$= \int_S E[Y|P(X) = p, CM = 1]f_1^S(p)dp - \int_S E[Y|P(X) = p, CM = 0]f_0^S(p)dp.$$

Analogously, to $f_1^S(p)$ in (2), $f_0^S(p)$ thereby is the density of the propensity scores $P = P(X)$ for people without case management ($CM = 0$) on the common support. By adding and subtracting the counterfactual outcome we obtain

$$\Delta_S = \underbrace{\int_S (E[Y|P(X) = p, CM = 1] - E[Y|P(X) = p, CM = 0])f_1^S(p)dp}_{\Delta \text{ in coefficients (impact case management)}} + \underbrace{\int_S E[Y|P(X) = p, CM = 0]\left(f_1^S(p) - f_0^S(p)\right)dp.}_{\Delta \text{ in characteristics}} \qquad (5)$$

Like in Eq. (4), the first term in (5) captures the difference that is due to differences in the coefficients and shows how the estimated average in the reintegration rate would change, if people with their predefined characteristics had case management. The second term is the difference that is due to differences in characteristics.

To understand the drivers of the overall effect, we additionally estimate separate heterogeneous effects for causes of incapacity where we observe more than 50 people receiving case management, gender, as well as age and income quartiles.

5 Results and robustness tests

5.1 Results

Table 2 below shows the result of the logistic regression, with "Cause: Circulation" as base cause category. The marginal effects shown are average marginal effects. For dummy variables we determine the average marginal effects for discrete variables. The case management variable is highly statistically significant with an average marginal effect of 0.068. This indicates that, ceteris paribus, the probability of reintegration increases by 6.8 percentage points if people with a disability participate in case management.

The results show that the probability for reintegration is significantly lower for males and elderly people. Very important is the time since notification of the disability. If notification takes place half a year after the event of

disability, the probability of reintegration decreases by 10.95 percentage points. This is in line with Sim (1999), who highlights the importance of intervening as soon as possible after a disabling event. Cancer significantly reduces the probability of reintegration. If a person has cancer the likelihood to return to work decreases by 13.5 percentage points. Compared to problems concerning the circulation as the cause for the disability, which is the base category, we find that the probability of reintegration is highest if the cause for the disability is an accident, followed by a pregnancy causing disability, addictions or if disability is caused by an illness related to the genitals. Lowest probability of reintegration is observed when the cause of disability is problems with the circulation, the base category, or digestion. For digestion the probability is not significantly different from circulation.

Table 3 below shows the result of the logistic regression when restricting the sample to the common support subpopulations.

The results show that, with exception of the cause "Accident", the significance does not change. The case management variable is still highly statistically significant, but now with an average marginal effect of 0.056, indicating that the probability of reintegration increases by 5.6 percentage points if people with a disability participate in case management.

For the common support subpopulation, we obtain an average reintegration rate for those who participate in case management of 51.7%, i.e. $E_S\left[Y^1|CM = 1\right] = 0.517$. The result from propensity sore matching indicates that the expected reintegration rate of those who participate in case management would only be 43.5%, if they did not participate in case management, that is $E_S\left[Y^0|CM = 1\right] = 0.435$. Thus, case management increases the reintegration rate for those who participate in case management by about 18.9% or 8.2 percentage points ($E\left[Y^1 - Y^0|CM = 1\right] = 0.082$). Bootstrapped standard errors, calculated based on $10'000$ bootstrap replications, provide evidence that this result is again statistically highly significant at the 1% level. The estimate of 8.2 percentage points is comparable to the impact of other training programs, e.g. Stephan and Pahnke (2011)

Table 2 Average marginal effects of the multivariate logistic regression

Variables	Logit: Reintegration	
	Marginal Effect	z-value
Case Management	0.068*** (0.022)	3.06
Male	−0.023*** (0.008)	−2.99
Age in years	−0.007*** (0.0003)	−21.34
Time in years since 01/01/2009	−0.043*** (0.002)	−27.18
Years since notification	−0.219*** (0.008)	−26.32
Income in 1000 CHF	0.0004*** (0.0001)	3.79
Cancer	−0.135*** (0.015)	−9.21
Cause: Digestion	0.038 (0.023)	1.63
Cause: Genitals	0.166*** (0.022)	7.42
Cause: Nervous	0.088*** (0.015)	5.97
Cause: Perception	0.072** (0.030)	2.45
Cause: Infections	0.154** (0.062)	2.49
Cause: Bones	0.178*** (0.014)	12.32
Cause: Metabolism	0.071** (0.030)	2.39
Cause: Accidents	0.258*** (0.015)	17.52
Cause: Addictions	0.212*** (0.049)	4.32
Cause: Pregnancy	0.224*** (0.035)	6.37
Cause: Others	0.062*** (0.020)	3.16
Number of Observations	18'237	
McFadden Pseudo R^2	0.127	

Standard errors in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Table 3 Average Marginal effects of the multivariate logistic regression with common support restrictions

Variables	Logit: Reintegration	
	Marginal effect	z-value
Case management	0.056** (0.023)	2.43
Male	−0.029*** (0.009)	−3.36
Age in years	−0.008*** (0.0004)	−19.87
Time in years since 01/01/2009	−0.040*** (0.002)	−22.43
Years since notification	−0.237*** (0.011)	−21.24
Income in CHF 1000	0.0005*** (0.0001)	4.33
Cancer	−0.124*** (0.016)	−7.87
Cause: Digestion	0.031 (0.024)	1.28
Cause: Genitals	0.162*** (0.025)	6.60
Cause: Nervous	0.085*** (0.015)	5.49
Cause: Perception	0.079** (0.031)	2.50
Cause: Infections	0.159** (0.065)	2.43
Cause: Bones	0.183*** (0.015)	11.88
Cause: Metabolism	0.067** (0.031)	2.17
Cause: Accidents	−0.029 (0.036)	−0.80
Cause: Addictions	0.219*** (0.052)	4.19
Cause: Pregnancy	0.215*** (0.038)	5.63
Cause: Others	0.075*** (0.021)	3.54
Number of Observations	14,487	
McFadden Pseudo R^2	0.099	

Standard errors in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

estimate the share in employment of participants in a provision of skills program for less than 4 months to be 9 percentage points higher at the end of the observation period than the comparison group of waiting persons.

The results of the Blinder-Oaxaca decomposition for logistic regressions are shown in Table 4. The magnitude of the characteristics effect is very small and, based on bootstrapped standard errors, statistically insignificant.

Based on our data, we can conclude that, on average, those who attend case management do not have significantly different characteristics than those who do not participate in case management and consequently the difference in the reintegration rates is not due to differences in characteristics.

Compared to this the difference that is due to the coefficients amounts to 7.6 percentage points. This part

Table 4 Blinder-Oaxaca decomposition for non-linear models

	Value	Proportion
Total difference	0.082*** (0.025)	100%
Difference due to characteristics	0.007 (0.008)	8.19%
Difference due to coefficients	0.076*** (0.025)	91.81%
Observations	18'237	

Bootstrapped standard errors, based on 10,000 replications, in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Table 5 Blinder-Oaxaca decomposition for non-linear models with common support restriction

	Value	Proportion
Total difference	0.096*** (0.025)	100%
Difference due to characteristics	0.036*** (0.008)	38.2%
Difference due to coefficients	0.060*** (0.025)	61.8%
Observations	14,487	

Bootstrapped standard errors, based on 10'000 replications, in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Table 6 Semiparametric decomposition of the reintegration rate for the common support

	Value	Proportion
Total difference	0.096*** (0.026)	100%
Difference due to characteristics	0.014*** (0.007)	14.6%
Difference due to coefficients	0.082*** (0.026)	85.4%
Observations	14'487	

Bootstrapped standard errors, based on 10'000 replications, in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

explains 91.8% of the total difference in reintegration rates between those who attend case management and those who do not and is highly statistically significant at the 1% level.

Table 5 shows the results of the Blinder-Oaxaca decomposition for logistic regressions when restricting the sample. For the common support subpopulation, the difference that is due to the coefficients amounts to 6.0 percentage points. For the common support subpopulation this part explains 61.8% of the total difference, which increases to 9.6 percentage points.

The results for the semiparametric decomposition are shown in Table 6 and indicate that our results obtained from the Blinder-Oaxaca decomposition are robust and not pending on the functional form assumption. The total gap in the average reintegration rates for the common support subpopulation between those who participate in case management and those who do not amounts to 9.6 percentage points. Only 1.4 percentage points are due to differences in characteristics. Compared to the Blinder-Oaxaca decomposition for logistic regressions we now find that the differences in characteristics for the common support subpopulation are statistically significant, even though this part is small in magnitude.

In line with the results obtained from the Blinder-Oaxaca decomposition for logistic regressions, the main part of the gap is due to differences in coefficients. The differences in coefficients explain 8.2 percentage points, or 85.4% of the total gap, and are again highly statistically significant. Therefore, we find strong evidence that the main part of the differences in the reintegration rates is due to case management itself and not driven by differences in the characteristics of the people in the group with case management and those of the group without case management.

To analyze the validity of propensity score matching we perform balancing tests. Therefore, we test whether the means in the characteristics differ significantly between the treatment and the control group. The results show

that, except for the variable "income", there are no statistical differences in the average values of the variables between the treated and control group members after propensity score weighting has been applied. Through propensity score weighting the average income in the control group is increased from CHF 69,517 to CHF 81,515, however, it is still significantly lower than the average income of CHF 88,492 in the treatment group. This implies that we can make the average incomes more comparable by applying propensity score weighting, but as there is still a significant difference, we cannot rule out a small bias due to the imbalanced income. Results are shown in Table 7.

To understand the drivers of the overall effect, we additionally estimate separate heterogeneous effects for causes of incapacity where we observe more than 50 people receiving case management, gender, as well as age and income quartiles. The results are shown in Tables 8, 9 and 10.

Table 7 Balancing test for the common support subpopulation after matching

Variable	Mean			Test st (1) = (2)
	Treated	**(2) Control weighted**	**(3) Control unweighted**	
Male	0.659	0.648	0.597	0.48
Age (in years)	41.949	41.677	47.552	0.52
Income (in CHF 1000)	88.492	81.515	69.517	2.47**
Time in years since 01/01/2009	6.403	6.411	4.964	− 0.09
Years since notification	0.397	0.403	0.575	− 0.41
Cancer	0.027	0.029	0.111	− 0.23
Cause: Circulation	0.067	0.064	0.085	0.27
Cause: Digestion	0.030	0.031	0.053	− 0.11
Cause: Genitals	0.015	0.016	0.054	− 0.19
Cause: Nervous	0.632	0.629	0.379	0.12
Cause: Perception	0.010	0.011	0.021	− 0.15
Cause: Infections	0.002	0.003	0.003	− 0.08
Cause: Bones	0.192	0.193	0.289	− 0.08
Cause: Metabolism	0.017	0.018	0.022	− 0.07
Cause: Accidents	0.002	0.002	0.014	0.16
Cause: Addictions	0.005	0.005	0.005	− 0.04
Cause: Pregnancy	0.012	0.014	0.012	− 0.20
Cause: Others	0.015	0.016	0.064	− 0.14

Table 8 Separate heterogeneous effects using propensity score matching for males and females as well as different causes with more than 50 observations in the case management group

	Cause "Nervous"	Cause "Bones"	Males only	Females only
Total difference	0.115*** (0.032)	− 0.071 (0.057)	0.066** (0.031)	0.151*** (0.042)
Difference due to characteristics	0.001 (0.009)	− 0.008 (0.020)	0.034*** (0.009)	0.002 (0.012)
Difference due to coefficients	0.114*** (0.032)	− 0.063 (0.061)	0.032 (0.032)	0.149*** (0.044)
Observations	5135	2957	8461	5384

Bootstrapped standard errors, based on 10,000 replications, in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

We find that the overall effect of case management is mainly driven by those with disability cause "Nervous". For this group the effect of case management is 11.4 percentage points and highly significant. Compared to this, the impact for those with cause "Bones" is insignificant. Moreover, we find that the effect is especially strong for females, where we estimate the impact of case management to be 14.9 percentage points.

Table 9 shows the results for different age quartiles. Case management is especially effective for people with an age between 39.36 and 49.29. Table 10 reports the results for different income quartiles. For all but the second quartile, the estimated parameter of the case management dummy variable is significant. The magnitude is largest for the first quartile where case management increases the probability for reintegration by 19.7 percentage points, followed by the fourth quartile where we estimate an increase of 9.9 percentage points in the probability for reintegration.

5.2 Robustness tests

Additionally, to the average marginal effects, we also estimate the marginal effects at the means. The results are similar to our baseline results and inferences do not change, see Table 11 in the Appendix. The impact of case management slightly increases to 8.1 percentage points.

Table 9 Separate heterogeneous effects using propensity score matching for different age quartiles

	Age Q1 Below 39.36	Age Q2 39.36 to 49.29	Age Q3 49.29 to 56.79	Age Q4 Above 56.79
Total difference	− 0.041 (0.041)	0.119*** (0.042)	0.042 (0.051)	0.219 (0.179)
Difference due to characteristics	− 0.051*** (0.012)	− 0.017 (0.016)	0.008 (0.018)	0.023 (0.025)
Difference due to coefficients	0.01 (0.041)	0.136*** (0.045)	0.034 (0.055)	0.195 (0.178)
Observations	3622	3622	3622	3621

Bootstrapped standard errors, based on 10'000 replications, in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Table 10 Separate heterogeneous effects using propensity score matching for different income quartiles

	Income Q1 Below 48.1	Income Q2 48.1 to 63.7	Income Q3 63.7 to 79.95	Income Q4 Above 79.95
Total difference	0.131 (0.086)	− 0.056 (0.054)	0.102** (0.048)	0.175*** (0.038)
Difference due to characteristics	− 0.066*** (0.025)	0.011 (0.018)	0.016 (0.015)	0.076*** (0.012)
Difference due to coefficients	0.197** (0.084)	− 0.067 (0.058)	0.087* (0.050)	0.099** (0.040)
Observations	3638	3561	3561	3619

Note: Some of the observations are on the threshold and therefore the number of observations is not equal in the different quartiles

Bootstrapped standard errors, based on 10'000 replications, in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

The descriptive statistics in Table 1 reveal, that for 71.9% of the insured people the cause of disability is either a problem with the nervous system (31.4%), bones (23.5%) or an accident (17.0%). To check, whether the impact is different for these people, we run the logistic regression including those three causes only, excluding insured persons whose cause of disability is not of one of these types. The effect of the causes "Nervous" and "Bones" changes due to the change of the base cause category to "Accidents". The gender dummy is insignificant in this specification. The estimated impact of case management on the reintegration into the labor force is 7.4 percentage points and highly statistically significant and, therefore, again like our baseline result of 6.8 percentage points. The results can be found in Table 12 in the Appendix.

Moreover, we re-do the propensity score matching using the optimal bandwidth based on cross-validation. Following Frölich (2004) we choose $0.01\sqrt{1.2}^{g-2}$ for $g = 1, \ldots, 59$ and ∞ as search grid for the bandwidth. The results do not change.

Insured people in our database who do not reintegrate into the labor force might pass away or enter retirement pension at one point in time. After this point in time reintegration is not possible any longer. In order to check whether they drive the results, we re-estimate the models by restricting the sample to insured persons who can potentially be reintegrated throughout a predefined period of time. Inferences again do not change, see Table 13 in the Appendix.

6 Conclusion

Using exclusive data from a Swiss life insurer we find that overall, 43.9% of the people that have been incapable to work could be reintegrated into the labor force market. The highest reintegration rates are found for people who are incapable to work due to pregnancy (69.2%) or an accident (59.9%). The lowest reintegration rates are seen for people who are incapable to work due to digestion (26.0%) or circulation (27.5%).

We examine the effectiveness of case management by analyzing the probability of labor force reintegration

of people with a disability. We estimate the probability using a logistic regression, controlling for several personal characteristics. Given average personal characteristics, we find that the estimated probability of reintegration is 6.8 percentage points higher if people participate in case management. When looking at the cause of the disability we find that the probability of reintegration is highest if the cause for the disability is an accident, followed by a pregnancy causing disability. Addictions or if disability is caused by an illness related to the genitals have the third and fourth highest probability of reintegration. Cancer significantly reduces the probability of reintegration by 13.5 percentage points.

Our results from propensity score matching, which does not rely on functional form assumptions, indicate that case management increases the reintegration rate for those who participate in case management by about 18.9% or 8.2 percentage points from 43.5 to 51.7%. This estimate is comparable to the impact of other training programs like the participation in a provision of skills program for less than 4 months (Stephan and Pahnke 2011).

Using the Blinder-Oaxaca decomposition for logistic regressions as well as the semiparametric decomposition based on propensity score matching we show that the majority of the difference in the reintegration rates between those who participate in case management and those who do not, is due to the case management program. Our estimates show that no more than 15% of the difference is due to differences in the average characteristics.

Strittmatter and Wunsch (2021) study different methods for analyzing the gender pay. Their results indicate that restricting the functional form how observed wage determinants and gender may affect wages can have a large impact on the estimated unexplained gender pay gap. They show that including wage determinants in a flexible way is crucial when using a parametric Blinder-Oaxaca decomposition. With a sufficiently large sample, however, they find that a flexible matching estimator is the better methodological choice, as matching does not restrict the heterogeneity in any way. Following the recommendations of Strittmatter and Wunsch (2021) our results based on matching are more reliable.

Our robustness tests show that our results are robust to various specifications. Based on our analysis, we recommend case management as an effective tool for the employees in order to have a higher probability of reintegrating to the labor market, governments to reduce social expenses, private insurers to reduce benefit payments and employers to retain important knowledge in the company.

It is acknowledged that the data does not allow us to account for several relevant variables. For example, we cannot observe the severity of the disability or the employment history of the people, which could possibly impact the estimates of the impact of case management. Another limitation is that the data does not allow us to follow Fredriksson and Johansson (2008) and apply a survival function matching estimator to estimate the time profile of case management. We cannot address the dynamic treatment selection problem as we cannot measure the elapsed duration in the group of people who do not participate in case management. Also unobserved individual circumstances might positively (e.g. the need for the money of the job income) or negatively (disability benefits are higher than unemployment benefits) impact the chance of reintegration. Another limitation is that we do not observe whether the comparison group members have obtained any other support during the time considered. This can potentially impact our estimates of the effectiveness of case management (Stephan 2008).

Our positive results may not be directly transferable to other types of case management, other providers, or countries. The effectiveness of other type of case management is not always clear in the literature. Even for the same cause, such as back-pain, some studies find positive effects (Jensen et al. 2012), while others find no effects (Oestergaard et al. 2020). In addition, there are country-specific differences. In the Netherlands it was found that better practice guidelines that promote an activating approach of the occupational physician to establish faster return-to-work does not result in an improved reintegration rate (van Beurden et al. 2017). However, in the Netherlands for a longer sick leave, each employee is already required to have a rehabilitation consultation. This is not the case for most other countries.

There are also some indisputable rules for successful case management in all facilities and countries. Interventions of the case manager should be multifaceted and specific to the individual needs. These range from optimizing the physical and mental health support (Wagner et al. 2016) to modified working arrangements (Waddell et al, 2008). In our study, case managers provide holistic support to the persons with disabilities. This ranges from support for the search for optimal medical therapies, legal issues, discussions with the employer, to the integration into the social network. By focusing on the specific needs of the person with a disability, we find a strong positive effect of case management on the reintegration rate. Future research should try to build on richer data sources that allow to include all relevant variables. Including these relevant variables will lead to a further improvement in understanding the effectiveness of case management. One interesting work would be to investigate which type of case management, depending on the cause of the incapacity to work, leads to the highest reintegration rate. From this analysis best practice recommendations for case managers could be developed. An improvement in this domain would be to the benefit of all parties involved.

Appendix

See Tables 11, 12, 13.

Table 11 Robustness check: Marginal effects at the mean

Variables	Logit: Reintegration	
	Marginal effect	**z-value**
Case Management	0.081*** (0.027)	3.04
Male	− 0.027*** (0.009)	− 2.98
Age in years	− 0.009*** (0.0004)	− 22.76
Time in years since 01/01/2009	− 0.051*** (0.002)	− 30.38
Years since notification	− 0.258*** (0.009)	− 29.35
Income in CHF 1000	0.0004*** (0.0001)	3.79
Cancer	− 0.153*** (0.016)	− 9.60
Cause: Digestion	0.045 (0.028)	1.62
Cause: Genitals	0.200*** (0.027)	7.33
Cause: Nervous	0.105*** (0.018)	5.85
Cause: Perception	0.086** (0.035)	2.43
Cause: Infections	0.183** (0.073)	2.50
Cause: Bones	0.213*** (0.018)	11.96
Cause: Metabolism	0.084** (0.036)	2.37
Cause: Accidents	0.301*** (0.017)	17.38
Cause: Addictions	0.253*** (0.057)	4.41
Cause: Pregnancy	0.266*** (0.041)	6.53
Cause: Others	0.074*** (0.024)	3.13
Number of Observations	18,237	
McFadden Pseudo R^2	0.127	

Standard errors in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Table 12 Robustness check: Logistic regression for causes nervous, bones and accident only using average marginal effects

Variables	Logit: Reintegration	
	Marginal effect	**z-value**
Case management	0.074*** (0.025)	3.00
Male	− 0.014 (0.009)	− 1.52
Age in years	− 0.007*** (0.0004)	− 17.10
Time in years since 01/01/2009	− 0.048*** (0.002)	− 24.56
Years since notification	− 0.232*** (0.010)	− 23.36
Income in 1000 CHF	0.0005*** (0.0001)	4.07
Cancer	− 0.252*** (0.031)	− 8.14
Cause: Nervous	− 0.172*** (0.010)	− 16.57
Cause: Bones	− 0.080*** (0.011)	− 7.42
Number of Observations	13,127	
McFadden Pseudo R^2	0.113	

Standard errors in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Table 13 Robustness check: Restricting sample to people who still are able to reintegrate using average marginal effects

Restriction	Coefficient on case management
Exclude all cases having been resolved by death or entry into retirement pension	0.064*** (0.022)
Exclude all cases having been resolved by entry into retirement pension	0.067*** (0.022)
Exclude all cases having been resolved by death	0.066*** (0.022)
Keep cases that are resolved either by death or by entry into retirement pension, but where the time between disability and retirement/death is at least 0.5 year	0.068*** (0.022)
Keep cases that are resolved either by death or by entry into retirement pension, but where the time between disability and retirement/death is at least 1 year	0.068*** (0.022)
Keep cases that are resolved either by death or by entry into retirement pension, but where the time between disability and retirement/death is at least 3 years	0.064*** (0.022)
Keep cases that are resolved either by death or by entry into retirement pension, but where the time between disability and retirement/death is at least 5 years	0.065*** (0.022)
Keep cases that are resolved either by death or by entry into retirement pension, but where the time between disability and retirement/death is at least 7 years	0.065*** (0.022)
Keep cases that are resolved either by death or by entry into retirement pension, but where the time between disability and retirement/death is at least 9 years	0.064*** (0.022)

Standard errors in parentheses

*** Indicates significance at the 1% level

** Indicates significance at the 5% level

* Indicates significance at the 10% level

Availability of data and materials
The data that support the findings of this study have been provided by AXA Life Switzerland, but the availability of these data is restricted as the data used for the current study are sensitive and not publicly available. Upon reasonable request, the authors will check for permission with AXA Life Switzerland.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] AXA Switzerland, General-Guisan-Strasse 42, 8400 Winterthur, Switzerland. [2] Furtwangen University, Robert-Gerwig-Platz 1, 78120 Furtwangen, Germany. [3] ZHAW School of Management and Law, St.-Georgen-Platz 2, 8401 Winterthur, Switzerland.

References

Abadie, A., Imbens, G.: Large sample properties of matching estimators for average treatment effects. Econometrica **74**(1), 235–267 (2006)

Bauer, T., Sinning, M.: An extension of the Blinder-Oaxaca decomposition to nonlinear models. AStA Adv. Stat> Anal. **92**(2), 197–206 (2008)

Blinder, A.: Wage discrimination: reduced form and structural estimates. J. Hum. Resour. **8**(4), 436–455 (1973)

Burkhauser, R.V., Daly, M.C., Ziebarth, N.R.: Protecting working-age people with disabilities: experiences of four industrialized nations. J. Labour Mark. Res. **49**(4), 367–386 (2016)

Busso, M., DiNardo, J., McCrary J. Finite sample properties of semiparametric estimators of average treatment effects. Unpublished manuscript, University of Michigan and University of Californa–Berkeley (2008).

Cardano, M., Costa, G., Demaria, M.: Social mobility and health in the Turin longitudinal study. Soc. Sci. Med. **58**(8), 1563–1574 (2004)

Clayton, S., Barr, B., Nylen, L., Burström, B., Thielen, K., Diderichsen, F., Dahl, E., Whitehead, M.: Effectiveness of return-to-work interventions for disabled people: a systematic review of government initiatives focused on changing the behaviour of employers. Eur. J. Public Health **22**(3), 434–439 (2012)

Crump, R.K., Hotz, V.J., Imbens, G.W., Mitnik, O.A.: Dealing with limited overlap in estimation of average treatment effects. Biometrika **96**(1), 187–199 (2009)

de Boer, A.G., Verbeek, J.H., Spelten, E.R., Uitterhoeve, A.L., Ansink, A.C., de Reijke, T.M., Kammeijer, M., Sprangers, M.A., van Dijk, F.J.: Work ability and return-to-work in cancer patients. Br. J. Cancer **98**(8), 1342–1347 (2008)

Dehejia, R., Wahba, S.: Causal effects in nonexperimental studies:reevaluating the evaluation of training programs. J. Am. Stat. Assoc. **94**(448), 1053–1062 (1999)

Ferlie, E., Montgomery, K., Pedersen, A. (eds.): The Oxford Handbook of health care management. Oxford University Press, Oxford (2016)

Fortin, N., Lemieux, T., Firpo, S.: Decomposition methods in economics. Handb. Labor Econ. **4**, 1–102 (2011)

Fredriksson, P., Johansson, P.: Dynamic treatment assignment: the consequences for evaluations using observational data. J. Bus. Econ. Stat. **26**(4), 435–445 (2008)

Frölich, M.: Finite-sample properties of propensity-score matching and weighting estimators. Rev. Econ. Stat. **86**(1), 77–90 (2004)

Frölich, M.: Propensity score matching without conditional independence assumption—with an application to the gender wage gap in the United Kingdom. Econ J. **10**(2), 359–407 (2007)

Galizzi, M., Leombruni, R., Pacelli, L.: Successful return to work during labor market liberalization: the case of Italian injured workers. J. Labour Mark. Res. **53**(1), 1–24 (2019)

Hahn, J.: On the role of the propensity score in efficient semiparametric estimation of average treatment effects. Econometrica **66**(2), 315–331 (1998)

Hansen E. (2006). Das Case / Care Management. In: Galuske M., Thole W. (eds) *Vom Fall zum Management*. VS Verlag für Sozialwissenschaften.

Heckman, J., Ichimura, H., Todd, P.: Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. Rev. Econ. Stud. **64**(4), 605–654 (1997)

Heckman, J., Ichimura, H., Todd, P.: Matching as an econometric evaluation estimator. Rev. Econ. Stud. **65**(2), 261–294 (1998)

Hofman, M., Ryan, J.L., Figueroa-Moseley, C.D., Pascal, J.P., Morrowa, G.R.: Cancer-Related Fatigue: The Scale of the Problem. Oncologist **12**, 4–10 (2007)

Imbens, G.W.: Nonparametric estimation of average treatment effects under exogeneity: a review. Rev. Econ. Stat. **86**(1), 4–29 (2004)

Jensen, L.D., Maribo, T., Schiøttz-Christensen, B., Madsen, F.H., Gonge, B., Christensen, M., Frost, P.: Counselling low-back-pain patients in secondary healthcare: a randomised trial addressing experienced workplace barriers and physical activity. Occup. Environ. Med. **69**(1), 21–28 (2012)

Lechner, M., Strittmatter, A.: Practical procedures to deal with common support problems in matching estimation. Economet. Rev. **38**(2), 193–207 (2019)

Liukko, J., Kuuva, N.: Cooperation of return-to-work professionals: the challenges of multi-actor work disability management. Disabil. Rehabil. **39**(15), 1466–1473 (2017)

Nadaraya, E.: On estimating regression. Theory of Probability and Its Applications **9**(1), 141–142 (1964)

Noh, J., Kim, J., Park, J., Kim, H., Kwon, Y.: Gender difference in relationship between health-related quality of life and work status. PLoS One **10**(12), e0143579 (2015)

Oaxaca, R.: Male-female wage differentials in urban labor markets. Int. Econ. Rev. **14**(3), 693–709 (1973)

Oestergaard, L.G., Christensen, F.B., Bünger, C.B., Søgaard, R., Holm, R., Helmig, P., Nielsen, C.: Does adding case management to standard rehabilitation affect functional ability, pain, or the rate of return to work after lumbar spinal fusion? A randomized controlled trial with two-year follow-up. Clin. Rehabil. **34**(3), 357–368 (2020)

Robins, J.M., Rotnitzky, A., Zhao, L.P.: Estimation of regression coefficients when some regressors are not always observed. J. Am. Stat. Assoc. **89**(427), 846–866 (1994)

Robroek, S., Schuring, M., Croezen, S., Stattin, M., Burdorf, A.: Poor health, unhealthy behaviors, and unfavorable work characteristics influence pathways of exit from paid employment among older workers in Europe: a four year follow-up study. Scand. J. Work Environ. Health **39**(2), 125–133 (2013)

Rosenbaum, P., Rubin, D.: The central role of the propensity score in observationalstudies for causal effects. Biometrika **70**(1), 41–55 (1983)

Rosenbaum, P., Rubin, D.: Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am. Stat. **39**(1), 33–38 (1985)

Rubin, D.: Estimating causal effects of treatments in randomized and nonrandomized studies. J. Educ. Psychol. **66**(5), 688–701 (1974)

Rubin, D.: Statistics and Causal Inference: Which Ifs Have Causal Answers. J. Am. Stat. Assoc. **81**(396), 961–962 (1986)

Schuring, M., Burdorf, L., Kunst, A., Mackenbach, J.: The effects of ill health on entering and maintaining paid employment: evidence in European countries. J. Epidemiol. Community Health **61**(7), 597–604 (2007)

Schuring, M., Mackenbach, J., Voorham, T., Burdorf, A.: The effect of re-employment on perceived health. J. Epidemiol. Community Health **65**(7), 639–644 (2011)

Segelken, H. R. Economists: How to slow the growth in disability claims. Cornell Chronicle (2014). Retrieved from https://news.cornell.edu/stories/2014/03/economists-how-slow-growth-disability-claims

Stephan, G.: The effects of active labor market programs in Germany: An investigation using different definitions of non-treatment. J. Econ. Stati. **228**(5–6), 586–611 (2008)

Stephan, G., Pahnke, A.: The relative effectiveness of selected active labor market programs: An empirical investigation for Germany. Manch. Sch. **79**(6), 1262–1293 (2011)

Strittmatter, A. and Wunsch, C. (2021). The Gender Pay Gap Revisited with Big Data: Do Methodological Choices Matter? *CESifo Working Paper* No. 8912.

Silverman, B.: Density estimation for statistics and data analysis. Chapman and Hall, London (1986)

Sim, J.: Improving return-to-work strategies in the United States disability programs, with analysis of program practices in Germany and Sweden. Soc. Secur. Bull. **62**(3), 41–50 (1999)

van den Berg, T., Elders, L., Burdorf, A.: Influence of health and work on early retirement. J. Occup. Environ. Med. **52**(6), 576–583 (2010)

van Beurden, K.M., Brouwers, E.P.M., Joosen, M.C.W., de Boer, M.R., van Weeghel, J., Terluin, B., van der Klink, J.J.L.: Effectiveness of an intervention to enhance occupational physicians' guideline adherence on sickness absence duration in workers with common mental disorders: a cluster-randomized controlled trial. J. Occup. Rehabil. **27**(4), 559–567 (2017)

Wacholder, S., McLaughlin, J.K., Silverman, D.T., Mandel, J.S.: Selection of controls in case-control studies: I Principles. Am. J. Epidemiol. **135**(9), 1019–1028 (1992a)

Wacholder, S., Silverman, D. T., McLaughlin, J. K. and Mandel, J. S. (1992b). Selection of controls in case-control studies: II. Types of controls. *American Journal of Epidemiology* 135(9), 1029–1041.

Waddell, G., Burton, K.: Is work good for your health and well-being? Stationery Office, London (2006)

Waddell, G., Burton, K., Kendall, N.A.: Vocational Rehabilitation – What Works, for Whom, and When? Report for the Vocational Rehabilitation Task Group. The Stationery Office, London (2008)

Wagner, S.L., Koehn, C., White, M.I., Harder, H.G., Schultz, I.Z., Williams-Whitt, K., Warje, O., Dionne, C.E., Koehoorn, M., Pasca, R., Hsu, V., McGuire, L., Schulz, W., Kube, D., Wright, M.D.: Mental Health interventions in the workplace and work outcomes: a best-evidence synthesis of systematic reviews. Int. J. Occup. Environ. Med. **7**(1), 1–14 (2016)

Watson, G.S.: Smooth regression analysis. Sankhyā Indian J. Stat. Series A **26**, 359–372 (1964)

Wheeler, P.M., Kearney, J.K., Harrison, C.A.: The US study of work incapacity and eintegration. Soc. Secur. Bull. **64**(1), 32–44 (2001)

Wooldridge, J.M.: Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge (2002)

Young, A.E., Roessler, R.T., Wasiak, R., McPherson, K.M., van Poppel, M.N., Anema, J.R.: A developmental conceptualization of return to work. J. Occup. Rehabil. **15**(4), 557–568 (2005)

Publisher's Note