BMJ Open

Evaluating Community-based Rehabilitation: can propensity score matching be applied to cross-sectional data?

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-022544
Article Type:	Research
Date Submitted by the Author:	22-Feb-2018
Complete List of Authors:	Mason, Catherine; LMU Munich, Department for Medical Information Processing, Biometry and Epidemiology (IBE) Sabariego, Carla; Ludwig-Maximilians University, Public Health and Health Services Research Thắng, Đoàn Mạnh ; Hoa Binh Department of Health Weber, Jörg; CBM eV
Keywords:	epidemiological methods, disability, social inequalities

Evaluating Community-based Rehabilitation: can propensity score matching be applied to cross-sectional data?

Authors: Catherine Mason[1*], Carla Sabariego[1], Đoàn Mạnh Thắng[2], Joerg Weber[3]

[1] Department for Medical Information Processing, Biometry and Epidemiology (IBE)
Ludwig-Maximilians-Universität München
Marchioninistraße 15
81377 Munich, Germany

Catherine.Mason@med.uni-muenchen.de
+49 (0)89 2180-78229

[2] Hoa Binh Department of Health (Sở Y tế tỉnh Hòa Bình)
Hoa Binh City, Vietnam

[3] CBM International
Bensheim, Germany

Keywords: disability, epidemiological methods, social inequalities

Word count: 3019

ABSTRACT

Objectives: Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalize opportunities and include people with disability in all aspects of life. The complexity of CBR and often limited resources lead to challenges when attempting to quantify its effectiveness, with randomization and longitudinal data rarely possible. Statistical methods, such as propensity score matching (PSM), offer an alternative approach to evaluate a treatment when randomization is not feasible. The aim of this study is to examine whether PSM can be an effective method to facilitate evaluations of results in CBR when data are cross-sectional.

Design: Cross-sectional survey

Setting and Participants: Data were collected using the World Health Organization's CBR Indicators in Vietnam, with treatment assignment (participating in CBR or not) determined by province of residence. 298 participants were selected through government records.

Results: PSM was conducted using one-to-one nearest neighbour method on ten covariates. Before matching, significant differences between groups were found for six of the ten covariates. PSM successfully adjusted for bias in all covariates in the matched sample (74 matched pairs). An independent t-test compared the outcome of "community inclusion" (a score based on selected indicators) between CBR and non-CBR participants both before and after matching, with CBR participants having significantly worse community inclusion scores than non-CBR participants. This result did not differ before and after matching.

Conclusion: PSM successfully reduced bias between groups, though this did not affect the tested outcome. PSM should be considered when analyzing cross-sectional CBR data, especially for international comparisons where differences between populations may be greater.

Strengths and limitations of this study

- The complexity of CBR and often limited resources available in the field lead to challenges in research attempting to quantify its effectiveness and to a heavy reliance on non-randomized cross-sectional data, implying the need for statistical approaches, such as PSM, to account for these limitations.
- PSM attempts to mimic randomization by creating a sample of participants who received the treatment (CBR participants) that is comparable on all observed covariates to participants who did not receive the treatment (non-CBR participants).
- The potential of using PSM for analyzing cross-sectional CBR data was demonstrated, as biases detected in the distribution of covariates between groups before matching were successfully eliminated.
- One of the main advantages of the CBR Indicators, namely the ability to use comparison individuals without disability from the community is lost; as PSM requires that all participants have a non-zero probability of receiving treatment meaning only people with disabilities can be included.
- PSM only controls for known covariates which means that there is a potential for bias if some covariates that affect the outcome are not included.

INTRODUCTION

Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalize opportunities and include people with disability in all aspects of community life. It is broadly defined as "a strategy within general community development for the rehabilitation, equalization of opportunities and social inclusion of all people with disabilities"[1]. The wide scope of CBR is further expanded through the various implementing stakeholders involved in CBR, including people with disabilities themselves, their families and communities, and the relevant governmental and non-governmental service sectors. It is due, at least in part, to this extensive definition that reliable and internationally comparable data to monitor and evaluate CBR are scarce. In an effort to synthesize global perspectives on CBR, the World Health Organization (WHO) developed their "Community-Based Rehabilitation Guidelines" in 2010 which have since become accepted as a conceptual framework for CBR[2]. With these guidelines, WHO emphasized the need for a common global framework for monitoring CBR in line with the Convention on the Rights of Persons with Disability (CRPD). With the launch of the global WHO CBR Indicators in 2015, there is now a standardized approach to do this[3,4].

The complexity of CBR leads to challenges in research attempting to quantify its effectiveness[5-7]. Fully experimental studies with randomization are rarely possible for both ethical and practical reasons, which inherently lead to limitations. The possibility of bias arises as the apparent difference in an outcome between two treatment groups may depend on characteristics that affected whether or not an individual received a given treatment, instead of being an actual effect of the treatment. For this reason there has been a recent emphasis on so-called natural experiments, where a range of primarily statistical approaches are used to evaluate a treatment or intervention when randomization is not feasible[8]. One such approach is propensity score matching (PSM).

PSM was first presented in 1983 by Rosenbaum and Rubin as a method to reduce bias due to confounding variables in observational studies[9]. It attempts to mimic randomization by creating a sample of participants who received the treatment that is comparable on all observed covariates to participants who did not receive the treatment. This effectively creates an experimental dataset where the comparison group is, on average, equivalent to individuals in the exposed group on all observed covariates[10-12]. A systematic review comparing 21 PSM studies to 63 RCTs on therapeutic interventions for acute coronary syndromes found that PSM produced more extreme treatment effect estimates when compared with those from RCTs, although these differences were rarely statistically significant[13]. A similar comparison including 20 propensity-score-based studies matched to RCT results was conducted examining critical care medicine and found that propensity-score-based studies report less beneficial effects of treatment in comparison to RCTs[14]. Despite some shortcomings, PSM provides a method for evaluating complex interventions where randomization is not possible.

PSM has been increasingly used in various research fields, including Public Health, to evaluate complex interventions[15]. CBR is considered a complex intervention, and data collection in the field is further hindered by low resources making quantitative longitudinal data collection infeasible and rarely done[6,7,16,17]. This implies that data analysis in the field of CBR relies heavily on cross-sectional data. PSM has already been successfully applied to cross-sectional data[18,19]. Therefore, the main objective of this paper is to examine whether PSM can be an effective method to facilitate evaluations of results in CBR when data are cross-sectional. Data used in the present study were

collected using the WHO CBR Indicators in Vietnam in 2016 with the assignment of persons to the treatment (CBR participants) and non-treatment group (non-CBR participants) determined by province of residence. PSM will be conducted on the outcome of *community inclusion* of people with disabilities, the ultimate goal of CBR in strong alinement with the CRPD, using a sum score of WHO CBR social indicators and an empowerment indicator.

METHODS

Data Collection

Data collection was conducted using the survey questionnaire accompanying the WHO CBR Indicators[3]. These indicators examine differences in health, education, social life, livelihood and empowerment between people with disabilities and other community members. There are two subsets of indicators: base indicators which are broad and should be used in all data collection activities to ensure comparability, and supplementary indicators which can provide more specific coverage, and can be selected depending on the specific CBR goals and strategies of a program. The indicators and corresponding questions used in this paper are presented in Table 1.

Data collection involved a multi-site cross-sectional survey in 2016 in two Vietnamese provinces: Huế, where CBR is fully implemented and all districts have CBR coverage through government implementation and through non-governmental organizations' (NGO) activities; and Hòa Bình, where CBR is not implemented by either government or NGOs. An Android mobile phone application (app), available from WHO for the CBR Indicators, was used to collect data during interviews (app free to download at: http://play.google.com/store/apps/details?id=com.universaltools.whocbrsurvey&hl=en).

People with disabilities were identified prior to the survey by government records. In both provinces a team of five local health care workers were trained by the lead researcher (CM) over two days on how to conduct interviews using the survey questions and the app. Data collection was supervised by CM. Data were collected anonymously and all respondents were informed of the purpose of the study, and then provided verbal (Huế) or written consent (Hòa Bình). In Huế the decision to provide verbal rather than written consent was justified since requiring written consent would embarrass illiterate participants, leading to a decreased willingness to answer further questions truthfully. In instances when the respondent had cognitive limitations that prevented them from being interviewed, or if the respondent was a minor, a proxy interview with a family member was performed. Ethical approval was obtained through the Ludwig-Maximilians-Universität Munich Ethics Commission.

Variables

Outcome Variable

To measure community inclusion, a sum score was created from the social base and supplementary questions, with the addition of the base question from empowerment. These questions all used the same response scale of *1(Not at all)* to *5(Completely)* with the final sum score ranging from 4 to 33, with higher scores indicating higher levels of inclusion (Table 1).

Table 1. WHO CBR Indicators and questions used to measure them. Base indicators are shown in bold. The response option for all questions ranged from 1 (Not at all) to 5 (Completely).

Component	Indicator	Survey Question
Social	**% of people with disability that feel valued as individuals by members of their community**	Do you feel that other people respect you? For example, do you feel that others value you as a person and listen to what you have to say?
	% of people with disability who make their own decisions about the personal assistance they need	Do you get to make decisions about the personal assistance that you need (who assists you, what type of assistance, when to get assistance)?
	% of people with disability make their own decisions about their personal relationships	Do you get to make your own decisions about your personal relationships, such as friends and family?
	% of people with disability who participate in artistic, cultural or religious activities	Do you get to participate in artistic, cultural or religious activities?
	% of people with disability who participate in mainstream recreational, leisure and sports activities	Do you get to participate in community recreational, leisure and sports activities?
	% of people with disability who know their legal rights	To what extent do you know your legal rights?
Empower-ment	**% of people with disability who make informed choices and decisions**	Do you get to make the big decisions in your life? For example, deciding who to live with, where to live, or how to spend your money?

Matching Variables

Data on *age* and *gender* were collected. Age was collected in categories (0-5, 6-12, 13-17, 18-24, 25-44, 45-64, and 65+) which were dichotomized for the analysis[20]. Though data on disability severity were not available, *general health status* was used as a proxy, using the question *"How would you rate your health today?"*[21]. A variable for *socio-economic status* (SES) was created using a sum score based on the questions "*What is the highest level of education you have achieved or are working to achieve?"* and "*Do you have enough money to meet your needs?".* The first question is commonly used in SES variable creation, and the second question targets wealth[22,23]. The variable *province of residence* corresponded to CBR coverage (no coverage in Hòa Bình, full CBR coverage in Huế). To account for economic differences between the provinces that might not be captured by SES, the covariate *receiving social protection* (such as for loss of income through old age, sickness or disability) was included. Covariates of *financial awareness* (knowing how to get financial services or social protection if needed), *having access to health services when needed,* and *having access to rehabilitation services when needed* were also included. A proxy for autonomy was captured through the covariates of *being involved in decision making regarding medical treatment* and *participating in a self-help group if desired*.

Missing Data

Missing data were low (2.25%). Multiple imputation (five imputations) using fully conditional specification (MICE package in R Studio Version 0.99.903) was used to replace missing data.

Analysis

Matching on the Propensity Score

The number of treated and untreated participants were similar (difference of n=4). Therefore, participants were matched using one-to-one nearest neighbour technique, which matched each treated unit to one control that was closest using calipers of width equal to 0.25 of the standard deviation (SD) of the logit of the estimated propensity score[24]. This implies that for a given treated participant, all the untreated participants are identified whose scores are within this specified

distance and then the best match is formed. If no match falls within this distance the participant is excluded. Participants were matched on ten covariates (see *Matching Variables*).

Balance Diagnostics

Baseline comparisons between the covariates were conducted before and after matching. Balance diagnosis was performed using the standardized difference method, which compares the difference in means of each covariate in units of the pooled standard deviation before and after matching[12]. Successful matching is indicated when the absolute standardized differences of means is less than 0.25[25].

Comparing Groups

For the community inclusion outcome, data matched on the ten covariates were compared using an independent t-test, as it cannot be assumed that the outcomes of matched individuals are correlated[26]. Bootstrapping was performed (1000 samples) in order to produce 95% confidence intervals (CI), which has been shown to account for uncertainty in the matching procedure[20].

A sensitivity analysis was performed using the Rosenbaum Bounds for Hodges-Lehmann Point Estimate to assess how robust the findings were to hidden bias due to unobserved covariates ('rbounds' package in R Studio Version 0.99.903). The maximum Gamma (the odds of differential assignment to treatment due to unobserved factors) was set to 2 with increments of 0.1 to test at which point the between group differences are no longer robust[27].

Data cleaning was performed using SPSS version 23 (copyright IBM Corporation). PSM was performed in R Studio (Version 0.99.903) using the 'MatchIt' package.

RESULTS

Data were available from 298 participants, of which 153(51.3%) were male, with a modal age group of 45-64(28.9%). The sample included 151 non-CBR participants and 147 CBR participants.

Before matching CBR participants had higher health status, were more likely to participate in a self-help group, more financially aware and more likely to be receiving social protection while they had worse access to rehabilitation services. Some age differences were also noted (Table 2). In the unmatched sample the absolute standardized difference across the 10 covariates ranged from 0.008 to 1.008 indicating bias.

When CBR participants were matched with non-CBR participants on the logit of the specified propensity score model, 74 matched pairs were formed. This meant that 49.7% of CBR participants were successfully matched to a control. PSM was successful in reducing bias between the covariates in the matched sample, as the standardized differences ranged from 0 to 0.147 with all values falling below the threshold value of 0.25[25] (Table 2).

Table 2. Baseline characteristics of CBR participants and non-CBR participants in the unmatched and matched samples. Absolute standardized differences of means are shown, with differences exceeding the threshold of 0.25 indicated in bold.

Variable	Unmatched Sample			Matched Sample		
	Mean	Mean	Std. dif. of	Mean	Mean	Std. dif. of

		No CBR (n=151)	With CBR (n=147)	means	No CBR (n=74)	With CBR (n=74)	means
Age	0-5	11 (7.2%)	6 (4.1%)	0.161	3 (4.1%)	5 (6.8%)	0.136
	6-12	19 (12.6%)	11 (7.5%)	0.193	7 (9.5%)	5 (6.8%)	0.102
	13-17	4 (2.6%)	6 (4.1%)	0.072	2 (2.7%)	1 (1.4%)	0.068
	18-24	12 (7.9%)	12 (8.2%)	0.008	7 (9.5%)	7 (9.5%)	0.000
	25-44	49 (32.5%)	32 (21.8%)	**0.258**	23 (31.1%)	22 (29.7%)	0.033
	45-64	42 (27.8%)	44 (29.9%)	0.046	21 (28.4%)	26 (35.1%)	0.147
	65+	14 (9.3%)	36 (24.5%)	**0.353**	11 (14.9%)	8 (10.8%)	0.094
Gender (male)		80 (53.0%)	73 (50.0%)	0.066	37 (50.0%)	42 (56.8%)	0.135
SES (range 1-10)		3.74±1.32	3.91±1.30	0.235	3.65±1.45	3.67±1.42	0.020
Health status (range 1-5)		2.89±0.77	3.37±0.70	**0.683**	3.05±0.75	3.14±0. 65	0.115
Receiving social protection		74 (49.0%)	117 (79.6%)	**1.008**	48 (64.9%)	52 (70.3%)	0.141
Access to health services		132 (87.4%)	126 (85.7%)	0.048	66 (89.2%)	66 (89.2%)	0.000
Access to rehabilitation services		128 (84.8%)	123 (83.7%)	**0.263**	29 (39.2%)	31 (41.9%)	0.054
Self-help group		63 (41.7%)	75 (51.0%)	**0.396**	31 (41.9%)	32 (43.2%)	0.027
Financial awareness		73 (48.3%)	122 (83.0%)	**0.789**	51 (68.9%)	55 (74.3%)	0.134
Involved in treatment decisions		47 (31.1%)	65 (44.2%)	0.137	65 (87.8%)	65 (87.8%)	0.000

Note: continuous variables are presented as means ± standard deviation; dichotomous variables are presented as N(%)

To test whether PSM affected the pre-defined outcome of community inclusion, the difference between groups before and after matching were assessed: similar significant differences were found. In the matched sample, CBR participants had worse community inclusion scores (mean=17.86, SD=6.30, 95%CI 16.33-19.24) than non-CBR participants (mean=20.93, SD=6.16, 95%CI 19.42-22.21); t(146)=2.996, p=0.003. The sensitivity analysis corroborated the results, showing that CBR participants had a median difference in community inclusion score 3.5 points lower than non-CBR participants (Gamma=0). When the Gamma value was increased to 2, the upper and lower bounds did not include zero indicating robust results[27]. These results did not differ from the results before PSM: community inclusion for participants with CBR (mean=18.61, SD=5.38) and without CBR (mean=20.64, SD=6.49); t(296)=2.935, p=0.004.

DISCUSSION

To our knowledge, this study presents the first use of PSM as a method for analyzing cross-sectional data in the field of CBR. The study analyzed data collected using the WHO CBR Indicators, and found that community inclusion scores of CBR participants were significantly lower than those of non-CBR participants after PSM. Despite bias being detected in the distribution of covariates between groups before matching, the results before PSM did not significantly differ from those after. We conclude that PSM can be successfully applied to cross-sectional CBR data, though in this case the bias reduction provided by PSM did not affect the tested outcome.

PSM has been applied only to longitudinal CBR data so far, but PSM studies using cross-sectional data are available from other fields. These studies had similar results in terms of the methodological success of PSM, but unlike our study they had final outcomes in line with their hypotheses. One such example is the study from Jalan and Ravallion which examines the effect of an employment-based poverty reduction program on income gain, accounting for pre-intervention and foregone income[19]. Through the trial of three PSM methods, they were able to reduce the differences

between the two populations and to demonstrate the effectiveness of the program. Another such example is the study from Becerril and Abdulai showing the positive impact of new maize farming technologies on per capita poverty outcomes[18]. Similarly to our study, they detected bias in the distribution of covariates between groups before matching, indicating that accounting for bias though PSM was important. In the field of CBR, PSM has been used to evaluate longitudinal CBR data in India, looking at livelihood and health outcomes[28,29]. PSM was used to reduce the bias between the CBR and non-CBR groups, with results showing that CBR participants had better health and livelihood outcomes, and that these differences generally increased over time. As in our study, these studies all showed bias between groups before matching which were reduced in the matched sample after PSM. However, none of these studies presented their outcome results before matching for comparison, so it cannot be determined if their final results were unaffected by matching as is the case in our study.

The results of the present study go against the anecdotal evidence that CBR has a positive influence on the lives of people with disabilities[6,7,30]. Results from longitudinal data indicate that CBR has a positive impact on receiving pensions, accessing paid jobs, accessing assistive devices, and personal-practical autonomy, with the impact increasing over time[28]. An explanation for our results could be that cross-sectional data do not allow causal inferences: results could simply point out that the province with highest problems has been selected for receiving CBR interventions. Additionally, this study focused on community inclusion - the ultimate goal of CBR - but when interpreting results it is also important to consider the specific targets of the program being examined. Though CBR aims to impact all aspects of the lives of people with disabilities, the program in Huế focuses specifically on improving the health of people with disabilities through physiotherapist visits and strengthening medical referral pathways. This could be a reason for the counter-intuitive results, and may demonstrate the importance of matching the indicators used with the targets of programs.

To our knowledge, this study is the first to implement the recently developed WHO CRB Indicators[4]. The study highlights how important it is to collect standardized data in the field of CBR in order to facilitate comparisons between groups and determine effectiveness of programs. One of the main advantages of the CBR Indicators and their data collection strategy is that they are easy to use in the field. The indicators allow for descriptive comparisons to be made easily, but in order for indicators to be used appropriately it is important to go beyond these descriptive results using inferential statistics. Furthermore, no single indicator or even a set of indicators is capable of capturing all changes in dynamic settings. The use of indicators alone has the potential limitation of collecting meaningless or misleading information,[31] and therefore they should be used as part of a broad evaluation strategy, in combination with qualitative and participatory evaluations[30]. Another way to reduce the limitations arising from indicator use is to continually test and re-assess the indicators[31]. In the case of the CBR Indicators, a priority should be to do this in partnership with communities in order to promote their uptake.

The use of PSM as a method for analysis of cross-sectional data collected from the CBR Indicators is conceptually strong, due to its ability to reduce bias due to confounding variables in observational studies[9]. However, the methodological limitations of PSM also need to be considered. PSM requires that each participant has a non-zero probability of receiving treatment, meaning only people with disabilities can be included in the analysis. Due to this, one of the main advantages of the CBR Indicators, namely the ability to use comparison individuals from the community, is lost[4]. Furthermore, PSM only controls for known covariates which means that there is a potential for bias if

some covariates that affect the outcome are not included[9]. For example, in this study no data were available on the ethnicity of participants, despite its known association with social disparities in Vietnam[32]. Another such covariate in this study could be disability severity, although this was partially adjusted for in both the participant selection, whereby all people with disabilities were identified using the same government disability criteria, and further in the analysis through the inclusion of the self-rated health covariate. Another limitation of PSM is that it leads to reduced sample size which could limit generalizability, though this is partly addressed through the provided sensitivity analysis. Further studies should look into additional statistical methods for analyzing the results obtained from the CBR Indicators.

Based on the present study, we recommend the further use and testing of the WHO CBR Indicators to increase standardized data collection in the field of CBR. In accompaniment to increased data collection, we recommend PSM as a method to reduce bias in cross-sectional CBR data analyses, especially for international comparisons where differences between populations may be greater than the within country differences observed in this study.

CONCLUSION

This study presents the first use of PSM as a method for analyzing cross-sectional CBR data. While randomized and longitudinal data are ideal for evaluations, this type of data collection is often not feasible in the field of CBR due to its high complexity and limited resources. The potential of using PSM for analyzing cross-sectional CBR data was demonstrated, though further research should investigate alternative inferential methods, such as cluster matching or adjusted regression, which may be more suitable in allowing for the comparison of the differences between persons with and without disabilities in line with the WHO CBR Indicators. We recommend that the questions and indicators be continually reviewed, and that future cross-sectional CBR studies use PSM to reduce bias when comparing groups.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING INTERESTS

The authors declare no competing interests.

DATA SHARING

The data are owned by the World Health Organization (WHO). Data are available from the WHO for researchers who meet the criteria for access to confidential data. Interested researchers can access the data by the same means the authors accessed them, by contacting WHO under disability@who.int. Statistical code is available from the corresponding author.

ACKNOWLEGEMENTS

The authors express gratitude to those from the Vietnamese Ministries and the local healthcare workers who lent their knowledge and field experience. Special thanks go out to the interviewers

who took the time to visit and speak with the participants; in Hue: Nguyen Thi Phung Diem, Nguyen Van Hong, Thuong Thi Huong Giang, Thian Cong Chirh, and Nguyen Thi Ngoc Anh and in Hoa Binh: Ha Thi Thoan, Vu Dury Hieu, Le Tleaal Hoa, Nguyen Quoc Dung, Le Vai Huy, and Nguyen Thanh.

AUTHOR CONTRIBUTIONS

Conceptualization: CM JW CS
Investigation: CM JW
Methodology: CM JW CS
Data curation: CM DMT
Formal analysis: CM
Project administration: CM JW CS
Resources: CM JW CS DMT
Supervision: CM DMT CS JW
Writing – original draft: CM CS JW
Writing – review & editing: CM JW CS DMT

REFERENCES

1 International Labour Organization, United Nations Educational, Scientific and Cultural Organization and the World Health Organization. CBR: A strategy for rehabilitation, equalization of opportunities, poverty reduction and social inclusion of people with disabilities: joint position paper. Geneva: World Health Organization, 2004.
2 World Health Organization, United Nations Educational, Scientific and Cultural Organization, & International Labour Organization. Community-based rehabilitation: CBR Guidelines. Geneva: World Health Organization, 2010.
3 World Health Organization, International Disability and Development Consortium. Capturing the difference we make- Community-based rehabilitation indicators manual. Geneva: World Health Organization, 2015.
4 Mason C, Weber J, Atasoy S, et al. Development of indicators for monitoring Community-Based Rehabilitation. *PLoS One* 2017;12(6):e0178418.
5 Finkenflügel H, Wolffers I, Huijsman R. The evidence base for community-based rehabilitation: a literature review. *Int J Rehabil Res* 2005;28(3):187-201.
6 Cornielje H, Velema JP, Finkenflugel H. Community based rehabilitation programmes: Monitoring and evaluation in order to measure results. *Lepr Rev* 2008;79(1):36-49.
7 Iemmi V, Blanchet K, Gibson LJ, et al. Community-based rehabilitation for people with physical and mental disabilities in low- and middle-income countries: a systematic review and meta-analysis. *Campbell Sys Rev* 2015;11(15).
8 Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health* 2012:jech-2011-200375.
9 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55.
10 Heckman JJ, Ichimura H, Todd P. Matching as an econometric evaluation estimator. *Rev Econ Stud* 1998;65(2):261-94.
11 Rosenbaum PR. The role of known effects in observational studies. *Biometrics* 1989;45(2):557-69.
12 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Research* 2011;46(3):399-424.

13 Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J* 2012;33(15):1893-901.
14 Zhang Z, Ni H, Xu X. Observational studies using propensity score analysis underestimated the effect sizes in critical care medicine. *J Clin Epidemiol* 2014;67(8):932-39.
15 Li M. Using the Propensity Score Method to Estimate Causal Effects A Review and Practical Guide. *Organ Res Methods* 2013;16(2):188-226.
16 Weber J, Polack S, Hartley S. An Online Survey on Identification of Evaluation Capacity, Needs and Current Practice of Programme Evaluation in Community-based Rehabilitation. *Disability CBR Inclusive Development* 2016;27(2):5-18.
17 Grandisson M, Hébert M, Thibeault R. A systematic review on how to conduct evaluations in community-based rehabilitation. *Disabil Rehabil* 2014;36(4):265-75.
18 Becerril J, Abdulai A. The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach. *World Devel* 2010;38(7):1024-35.
19 Jalan J, Ravallion M. Estimating the benefit incidence of an antipoverty program by propensity-score matching. *J Bus Econ Stat* 2003;21(1):19-30.
20 Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci* 2010;25(1):1.
21 Wu S, Wang R, Zhao Y, et al. The relationship between self-rated health and objective health status: a population-based study. *BMC Public Health* 2013;13(1):320.
22 Pollack CE, Chideya S, Cubbin C, et al. Should health studies measure wealth?: a systematic review. *Am J Prev Med* 2007;33(3):250-64.
23 Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc* 2007;99(9):1013.
24 Cochran WG, Rubin DB. Controlling bias in observational studies: A review. *Sankhya Ser A* 1973;417-46.
25 Ho DE, Imai K, King G, et al. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal* 2007;15(3):199-236.
26 Schafer JL, Kang J. Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychol Methods* 2008;13(4):279.
27 Keele L. An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. Columbus, OH: White Paper, 2010.
28 Mauro V, Biggeri M, Deepak S, et al. The effectiveness of community-based rehabilitation programmes: an impact evaluation of a quasi-randomised trial. *J Epidemiol Community Health* 2014;68(11):1102-08.
29 Biggeri M, Deepak S, Mauro V, et al. Do community-based rehabilitation programmes promote the participation of persons with disabilities? A case control study from Mandya District, in India. *Disabil Rehabil* 2014;36(18):1508-17.
30 Weber J, Grech S, Polack S. Towards a 'mind map' for evaluative thinking in Community Based Rehabilitation: reflections and learning. *Disability Global South* 2016;3(2):951-79.
31 Implementing the vision: Addressing challenges to results-focused management and budgeting. Implementation Challenges in Results Focused Management and Budgeting; 2002 11-12 February; Paris. OECD.
32 Taylor P. Social inequality in Vietnam and the challenges to reform. Singapore: Institute of Southeast Asian Studies 2004:23,212.

BMJ Open

Can propensity score matching be applied to cross-sectional data to evaluate Community-based Rehabilitation? Results of a survey implementing the World Health Organization's Community-based Rehabilitation Indicators in Vietnam

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-022544.R1
Article Type:	Research
Date Submitted by the Author:	12-Jul-2018
Complete List of Authors:	Mason, Catherine; LMU Munich, Department for Medical Information Processing, Biometry and Epidemiology (IBE) Sabariego, Carla; Ludwig-Maximilians University, Public Health and Health Services Research Thắng, Đoàn Mạnh ; Hoa Binh Department of Health Weber, Jörg; CBM eV
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Global health, Public health, Epidemiology
Keywords:	epidemiological methods, disability, social inequalities

Can propensity score matching be applied to cross-sectional data to evaluate Community-based Rehabilitation? Results of a survey implementing the World Health Organization's Community-based Rehabilitation Indicators in Vietnam

Authors: Catherine Mason[1*], Carla Sabariego[1], Đoàn Mạnh Thắng[2], Joerg Weber[3]

[1] Department for Medical Information Processing, Biometry and Epidemiology (IBE)
Ludwig-Maximilians-Universität München
Marchioninistraße 15
81377 Munich, Germany

Catherine.Mason@med.uni-muenchen.de
+49 (0)89 2180-78229

[2] Hoa Binh Department of Health (Sở Y tế tỉnh Hòa Bình)
Hoa Binh City, Vietnam

[3] CBM International
Bensheim, Germany

Keywords: disability, epidemiological methods, social inequalities

Word count: 4287

ABSTRACT

Objectives: Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalize opportunities and include people with disabilities in all aspects of life. The complexity of CBR and often limited resources lead to challenges when attempting to quantify its effectiveness, with randomization and longitudinal data rarely possible. Statistical methods, such as propensity score matching (PSM), offer an alternative approach to evaluate a treatment when randomization is not feasible. The aim of this study is to examine whether PSM can be an effective method to facilitate evaluations of results in CBR when data are cross-sectional.

Design: Cross-sectional survey

Setting and Participants: Data were collected using the World Health Organization's CBR Indicators in Vietnam, with treatment assignment (participating in CBR or not) determined by province of residence. 298 participants were selected through government records.

Results: PSM was conducted using one-to-one nearest neighbour method on ten covariates. Before matching, significant differences between groups were found for six of the ten covariates. PSM successfully adjusted for bias in all covariates in the matched sample (74 matched pairs). A paired t-test compared the outcome of "community inclusion" (a score based on selected indicators) between CBR and non-CBR participants both before and after matching, with CBR participants found to have significantly worse community inclusion scores (mean=17.86, SD=6.30, 95%CI 16.45-19.32) than non-CBR participants (mean=20.93, SD=6.16, 95%CI 19.50-22.35); t(73)=3.068, p=0.001. This result did not differ before and after matching.

Conclusion: PSM successfully reduced bias between groups, though its application did not affect the tested outcome. PSM should be considered when analyzing cross-sectional CBR data, especially for international comparisons where differences between populations may be greater.

Strengths and limitations of this study

- The complexity of CBR and often limited resources available in the field lead to challenges in research attempting to quantify its effectiveness and to a heavy reliance on non-randomized cross-sectional data, implying the need for statistical approaches, such as PSM, to account for these limitations.
- PSM attempts to mimic randomization by creating a sample of participants who received the treatment (CBR participants) that is comparable on all observed covariates to participants who did not receive the treatment (non-CBR participants).
- The potential of using PSM for analyzing cross-sectional CBR data was demonstrated, as biases detected in the distribution of covariates between groups before matching were successfully eliminated.
- One of the main advantages of the CBR Indicators, namely the ability to use comparison individuals without disability from the community is lost; as PSM requires that all participants have a non-zero probability of receiving treatment meaning only people with disabilities can be included.

- PSM only controls for known covariates which means that there is a potential for bias if some covariates that affect the outcome are not included.

INTRODUCTION

Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalize opportunities and include people with disabilities in all aspects of community life. It is broadly defined as "a strategy within general community development for the rehabilitation, equalization of opportunities and social inclusion of all people with disabilities"[1]. The wide scope of CBR is further expanded through the various implementing stakeholders involved in CBR, including people with disabilities themselves, their families and communities, and the relevant governmental and non-governmental service sectors. It is due, at least in part, to this extensive definition that reliable and internationally comparable data to monitor and evaluate CBR are scarce. In an effort to synthesize global perspectives on CBR, the World Health Organization (WHO) developed their "Community-Based Rehabilitation Guidelines" in 2010 which have since become accepted as a conceptual framework for CBR[2]. With these guidelines, WHO emphasized the need for a common global framework for monitoring CBR in line with the Convention on the Rights of Persons with Disability (CRPD). With the launch of the global WHO CBR Indicators in 2015, there is now a standardized approach to do this[3,4].

The complexity of CBR leads to challenges in research attempting to quantify its effectiveness[5-7]. Fully experimental studies with randomization are rarely possible for both ethical and practical reasons, which inherently lead to limitations. The possibility of bias arises as the apparent difference in an outcome between two treatment groups may depend on characteristics that affected whether or not an individual received a given treatment, instead of being an actual effect of the treatment. For this reason there has been a recent emphasis on so-called natural experiments, where a range of primarily statistical approaches are used to evaluate a treatment or intervention when randomization is not feasible[8]. One such approach is propensity score matching (PSM).

PSM was first presented in 1983 by Rosenbaum and Rubin as a method to reduce bias due to confounding variables in observational studies[9]. It attempts to mimic randomization by creating a sample of participants who received the treatment that is comparable on all observed covariates to participants who did not receive the treatment. This effectively creates an experimental dataset where the comparison group is, on average, equivalent to individuals in the exposed group on all observed covariates[10-12]. A systematic review comparing 21 PSM studies to 63 RCTs on therapeutic interventions for acute coronary syndromes found that PSM produced more extreme treatment effect estimates when compared with those from RCTs, although these differences were rarely statistically significant[13]. A similar comparison including 20 propensity-score-based studies matched to RCT results was conducted examining critical care medicine and found that propensity-score-based studies report less beneficial effects of treatment in comparison to RCTs[14]. Despite some shortcomings, PSM provides a method for evaluating complex interventions where randomization is not possible.

PSM has been increasingly used in various research fields, including Public Health, to evaluate complex interventions[15]. CBR is considered a complex intervention, and data collection in the field is further hindered by low resources making quantitative longitudinal data collection infeasible and rarely done[6,7,16,17]. This implies that data analysis in the field of CBR relies heavily on cross-sectional data. PSM has already been successfully applied to cross-sectional data[18,19]. Therefore, the main objective of this paper is to examine whether PSM can be an effective method to facilitate evaluations of results in CBR when data are cross-sectional. Data used in the present study were

collected using the WHO CBR Indicators in Vietnam in 2016 with the assignment of persons to the treatment (CBR participants) and non-treatment group (non-CBR participants) determined by province of residence. PSM will be conducted on the outcome of *community inclusion* of people with disabilities, the ultimate goal of CBR in strong alinement with the CRPD, using a sum score of WHO CBR social indicators and an empowerment indicator.

METHODS

Data Collection

Data collection was conducted using the survey questionnaire accompanying the WHO CBR Indicators[3]. These indicators examine differences in health, education, social life, livelihood and empowerment between people with disabilities and other community members. There are two subsets of indicators: base indicators which are broad and should be used in all data collection activities to ensure comparability, and supplementary indicators which can provide more specific coverage, and can be selected depending on the specific CBR goals and strategies of a program. The indicators and corresponding questions used in this paper are presented in table 1.

Data collection involved a multi-site cross-sectional survey in 2016 in two Vietnamese provinces: Huế, where CBR is fully implemented and all districts have CBR coverage through government implementation and through non-governmental organizations' (NGO) activities; and Hòa Bình, where CBR is not implemented by either government or NGOs. The Huế CBR program began in 2009 in cooperation with the Huế Rehabilitation Hospital. The program focused mainly on activities to increase capacity building for CBR workers, not only in terms of rehabilitation skills, but also working to improve their counselling and networking skills. The other focus of the program was to strengthen referral pathways for people with disabilities so that they could be connected with other existing services in the province, such as schools with teachers who were trained to support students with disabilities and vocational training centers. An Android mobile phone application (app), available from WHO for the CBR Indicators, was used to collect data during interviews (app free to download at:
http://play.google.com/store/apps/details?id=com.universaltools.whocbrsurvey&hl=en).

People with disabilities were identified prior to the survey by government records. In both provinces a team of five local health care workers were trained by the lead researcher (CM) over two days on how to conduct interviews using the survey questions and the app. Data collection was supervised by CM. Data were collected during face-to-face interviews with data recorded anonymously. All respondents were informed of the purpose of the study, and then provided verbal (Huế) or written consent (Hòa Bình). In Huế the decision to provide verbal rather than written consent was justified since requiring written consent would embarrass illiterate participants, leading to a decreased willingness to answer further questions truthfully. In instances when the respondent had cognitive limitations that prevented them from being interviewed, or if the respondent was a minor, a proxy interview with a family member was performed. Ethical approval was obtained through the Ludwig-Maximilians-Universität Munich Ethics Commission and by the local provincial Ministries of Health.

Variables

Outcome Variable

To measure community inclusion, a sum score was created from the social base and supplementary questions, with the addition of the base question from empowerment. These questions all used the same response scale of *1(Not at all)* to *5(Completely)* with the final sum score ranging from 4 to 33, with higher scores indicating higher levels of inclusion (Table 1).

Table 1. WHO CBR Indicators and questions used to measure them. Base indicators are shown in bold. The response option for all questions ranged from 1 (Not at all) to 5 (Completely).

Component	Indicator	Survey Question
Social	**% of people with disability that feel valued as individuals by members of their community**	Do you feel that other people respect you? For example, do you feel that others value you as a person and listen to what you have to say?
	% of people with disability who make their own decisions about the personal assistance they need	Do you get to make decisions about the personal assistance that you need (who assists you, what type of assistance, when to get assistance)?
	% of people with disability make their own decisions about their personal relationships	Do you get to make your own decisions about your personal relationships, such as friends and family?
	% of people with disability who participate in artistic, cultural or religious activities	Do you get to participate in artistic, cultural or religious activities?
	% of people with disability who participate in mainstream recreational, leisure and sports activities	Do you get to participate in community recreational, leisure and sports activities?
	% of people with disability who know their legal rights	To what extent do you know your legal rights?
Empower-ment	**% of people with disability who make informed choices and decisions**	Do you get to make the big decisions in your life? For example, deciding who to live with, where to live, or how to spend your money?

Matching Variables

Matching variables were those available from the WHO CBR Indicators, and were selected based on their theoretical association with community inclusion, primarily using CBR Guidelines[2]. Data on *age* and *gender* were collected. Age was collected in categories (0-5, 6-12, 13-17, 18-24, 25-44, 45-64, and 65+) which were dichotomized for the analysis[20]. Though data on disability severity were not available, *general health status* was used as a proxy, using the question *"How would you rate your health today?"*[21]. A variable for *socio-economic status* (SES) was created using a sum score based on the questions "*What is the highest level of education you have achieved or are working to achieve?"* and "*Do you have enough money to meet your needs?".* The first question is commonly used in SES variable creation, and the second question targets wealth[22,23]. The variable *province of residence* corresponded to CBR coverage (no coverage in Hòa Bình, full CBR coverage in Huế). To account for economic differences between the provinces that might not be captured by SES, the covariate *receiving social protection* (such as for loss of income through old age, sickness or disability) was included. Covariates of *financial awareness* (knowing how to get financial services or social protection if needed), *having access to health services when needed,* and *having access to rehabilitation services when needed* were also included. A proxy for autonomy was captured through the covariates of *being involved in decision making regarding medical treatment* and *participating in a self-help group if desired* (see Supplementary Table). Seeing as the CBR program in Huế focused on increasing referral pathways within the medical and education sectors, the questions derived from the education component and many from the medical component were not included as matching variables, since including covariates associated with CBR participation but not with community inclusion decrease model precision[24].

Missing Data

Missing data were low (2.25%). Multiple imputation (five imputations) using fully conditional specification (MICE package [25] in R Studio Version 0.99.903) was used to replace missing data.

Analysis

Matching on the Propensity Score

The number of treated and untreated participants were similar (difference of n=4). Therefore, participants were matched using one-to-one nearest neighbour technique, which matched each treated unit to one control that was closest using calipers of width equal to 0.25 of the standard deviation (SD) of the logit of the estimated propensity score without iteration[26]. This implies that for a given treated participant, all the untreated participants are identified whose scores are within this specified distance and then the best match is formed. If no match falls within this distance the participant is excluded. Participants were matched on ten covariates (see *Matching Variables*).

Balance Diagnostics

Baseline comparisons between the covariates were conducted before and after matching. Balance diagnosis was performed using the standardized difference method, which compares the difference in means of each covariate in units of the pooled standard deviation before and after matching[12]. Successful matching is indicated when the absolute standardized differences of means is less than 0.25[27].

Comparing Groups

For the community inclusion outcome, data matched on the ten covariates were compared using a paired t-test[28]. Bootstrapping was performed (1000 samples) in order to produce 95% confidence intervals (CI), which has been shown to account for uncertainty in the matching procedure[20].

A sensitivity analysis was performed using the Rosenbaum Bounds for Hodges-Lehmann Point Estimate to assess how robust the findings were to hidden bias due to unobserved covariates ('rbounds' package [29] in R Studio Version 0.99.903). The maximum Gamma (the odds of differential assignment to treatment due to unobserved factors) was set to 2 with increments of 0.1 to test at which point the between group differences are no longer robust[29].

Data cleaning was performed using SPSS version 23 (copyright IBM Corporation). PSM was performed in R Studio (Version 0.99.903) using the 'MatchIt' package[30].

Patient and Public Involvement

Participants were not directly involved in the development of the research question, study design, recruitment or conduct of the study. However, in the province of Huế (where CBR is implemented) participants are continually involved in the development of the CBR program, as CBR is participatory in nature. It was through their motivation – stemming from the need to prove to the national government and international donors that their intervention has an impact in order to receive funds – that the survey was conducted in the first place. A study report was submitted to the Huế and Hòa

Bình Ministries of Health which presented simple numeric and graphic descriptive findings which were to be communicated to participants.

RESULTS

Data were available from 298 participants. In Huế, 575 people with disabilities were identified by government records and 147 were included, while in Hòa Bình 375 people were identified by government records and 151 were included (sample size calculated using an alpha significance level of 0.05 and power of 90%). Included participants were randomly selected from the complete list. After the random selection, each interviewer was assigned a group of selected participants based on their geographic location. Of the randomly selected participants, one in Hòa Bình could not be contacted so another participant was selected. In both provinces none of the invited participants refused participation. Males comprised 153(51.3%) of the participants, with a modal age group of 45-64(28.9%) (see table 2 for further descriptives).

Before matching CBR participants had higher health status, were more likely to participate in a self-help group, more financially aware and more likely to be receiving social protection while they had worse access to rehabilitation services. Some age differences were also noted (Table 2). In the unmatched sample the absolute standardized difference across the 10 covariates ranged from 0.008 to 1.008 indicating bias.

When CBR participants were matched with non-CBR participants on the logit of the specified propensity score model, 74 matched pairs were formed. This meant that 49.7% of CBR participants were successfully matched to a control. PSM was successful in reducing bias between the covariates in the matched sample, as the standardized differences ranged from 0 to 0.147 with all values falling below the threshold value of 0.25[27] (Table 2).

Table 2. Baseline characteristics of CBR participants and non-CBR participants in the unmatched and matched samples. Absolute standardized differences of means are shown, with differences exceeding the threshold of 0.25 indicated in bold.

Variable		Unmatched Sample			Matched Sample		
		Mean No CBR (n=151)	Mean With CBR (n=147)	Std. dif. of means	Mean No CBR (n=74)	Mean With CBR (n=74)	Std. dif. of means
Age	0-5	11 (7.2%)	6 (4.1%)	0.161	3 (4.1%)	5 (6.8%)	0.136
	6-12	19 (12.6%)	11 (7.5%)	0.193	7 (9.5%)	5 (6.8%)	0.102
	13-17	4 (2.6%)	6 (4.1%)	0.072	2 (2.7%)	1 (1.4%)	0.068
	18-24	12 (7.9%)	12 (8.2%)	0.008	7 (9.5%)	7 (9.5%)	0.000
	25-44	49 (32.5%)	32 (21.8%)	**0.258**	23 (31.1%)	22 (29.7%)	0.033
	45-64	42 (27.8%)	44 (29.9%)	0.046	21 (28.4%)	26 (35.1%)	0.147
	65+	14 (9.3%)	36 (24.5%)	**0.353**	11 (14.9%)	8 (10.8%)	0.094
Gender (male)		80 (53.0%)	73 (50.0%)	0.066	37 (50.0%)	42 (56.8%)	0.135
SES (range 1-10)		3.74±1.32	3.91±1.30	0.235	3.65±1.45	3.67±1.42	0.020
Health status (range 1-5)		2.89±0.77	3.37±0.70	**0.683**	3.05±0.75	3.14±0. 65	0.115
Receiving social protection		74 (49.0%)	117 (79.6%)	**1.008**	48 (64.9%)	52 (70.3%)	0.141
Access to health services		132 (87.4%)	126 (85.7%)	0.048	66 (89.2%)	66 (89.2%)	0.000
Access to rehabilitation services		128 (84.8%)	123 (83.7%)	**0.263**	29 (39.2%)	31 (41.9%)	0.054
Self-help group		63 (41.7%)	75 (51.0%)	**0.396**	31 (41.9%)	32 (43.2%)	0.027
Financial awareness		73 (48.3%)	122 (83.0%)	**0.789**	51 (68.9%)	55 (74.3%)	0.134
Involved in treatment		47 (31.1%)	65 (44.2%)	0.137	65 (87.8%)	65 (87.8%)	0.000

decisions						

Note: continuous variables are presented as means ± standard deviation; dichotomous variables are presented as N(%)

To test whether PSM affected the pre-defined outcome of community inclusion, the difference between groups before and after matching were assessed: similar significant differences were found. In the matched sample, CBR participants had worse community inclusion scores (mean=17.86, SD=6.30, 95%CI 16.45-19.32) than non-CBR participants (mean=20.93, SD=6.16, 95%CI 19.50-22.35); t(73)=3.068, p=0.001. The sensitivity analysis corroborated the results, showing that CBR participants had a median difference in community inclusion score 3.5 points lower than non-CBR participants (Gamma=0). When the Gamma value was increased to 2, the upper and lower bounds did not include zero indicating robust results[29]. In a further sensitivity analysis, to ensure that the covariate of "access to rehabilitation" did not bias the model by being more strongly associated with receiving CBR rather than with the outcome of community inclusion, the model was run excluding this variable. The new model resulted in 75 matched pairs with all standardized differences falling below the threshold. The results of the t-test did not differ from the model including access to rehabilitation; CBR participants had worse community inclusion scores (mean=18.11, SD=5.981, 95%CI 16.72-19.47) than non-CBR participants (mean=21.17, SD=6.381, 95%CI 19.67-22.60); t(74)=3.310, p=0.0014.

Overall, the results did not differ from the results before PSM: community inclusion for participants with CBR (mean=18.61, SD=5.38) and without CBR (mean=20.64, SD=6.49); t(296)=2.935, p=0.004 using an independent t-test.

DISCUSSION

To our knowledge, this study presents the first use of PSM as a method for analyzing cross-sectional data in the field of CBR. The study analyzed data collected using the WHO CBR Indicators, and found that community inclusion scores of CBR participants were significantly lower than those of non-CBR participants after PSM. Despite bias being detected in the distribution of covariates between groups before matching, the results before PSM did not significantly differ from those after. We conclude that PSM can be successfully applied to cross-sectional CBR data, though in this case the bias reduction provided by PSM did not affect the tested outcome.

PSM has been applied only to longitudinal CBR data so far, but PSM studies using cross-sectional data are available from other fields. These studies had similar results in terms of the methodological success of PSM, but unlike our study they had final outcomes in line with their hypotheses. One such example is the study from Jalan and Ravallion which examines the effect of an employment-based poverty reduction program on income gain, accounting for pre-intervention and foregone income[19]. Through the trial of three PSM methods, they were able to reduce the differences between the two populations and to demonstrate the effectiveness of the program. Another such example is the study from Becerril and Abdulai showing the positive impact of new maize farming technologies on per capita poverty outcomes[18]. Similarly to our study, they detected bias in the distribution of covariates between groups before matching, indicating that accounting for bias though PSM was important. In the field of CBR, PSM has been used to evaluate longitudinal CBR data in India, looking at livelihood and health outcomes[31,32]. PSM was used to reduce the bias between the CBR and non-CBR groups, with results showing that CBR participants had better health and

livelihood outcomes, and that these differences generally increased over time at both four years and seven years. In our study, data was collected seven years after the program began, which would make the timing comparable and it is therefore plausible that the effect of CBR in our study could already be quantifiable. As in our study, these studies all showed bias between groups before matching which were reduced in the matched sample after PSM. However, none of these studies presented their outcome results before matching for comparison, so it cannot be determined if their final results were unaffected by matching as is the case in our study.

The results of the present study go against the anecdotal evidence that CBR has a positive influence on the lives of people with disabilities[6,7,33]. Results from longitudinal data indicate that CBR has a positive impact on receiving pensions, accessing paid jobs, accessing assistive devices, and personal-practical autonomy, with the impact increasing over time[31]. An explanation for our results could be that cross-sectional data do not allow causal inferences: results could simply point out that the province with highest problems has been selected for receiving CBR interventions. While the cross-sectional data collected in this study represent the first quantitative data from the region and therefore an important foundation for future work, the results emphasizes the general need for further collection and publication of CBR data, especially longitudinal data. Additionally, this study focused on community inclusion - the ultimate goal of CBR - but when interpreting results it is also important to consider the specific targets of the program being examined. Though CBR aims to impact all aspects of the lives of people with disabilities to increase community inclusion, the program in Huế does not directly target community inclusion. The program focuses on increasing the capacity of CBR workers and on strengthening referral pathways with the medical and educational sectors. Through these activities the community inclusion of people with disabilities should improve over time, but since community inclusion was not the direct target of the program, the community inclusion effects might only appear after a longer period, which could be a reason for the counter-intuitive results. Therefore, when assessing a program in its early stages, it may be more important to match the indicators used with the specific targets of programs.

To our knowledge, this study is the first to implement the recently developed WHO CRB Indicators[4]. The study highlights how important it is to collect standardized data in the field of CBR in order to facilitate comparisons between groups and determine effectiveness of programs. One of the main advantages of the CBR Indicators and their data collection strategy is that they are easy to use in the field. The indicators allow for descriptive comparisons to be made easily, but for indicators to be used appropriately it is important to go beyond these descriptive results using inferential statistics. Furthermore, no single indicator or even a set of indicators is capable of capturing all changes in dynamic settings. The use of indicators alone has the potential limitation of collecting meaningless or misleading information,[34] and therefore they should be used as part of a broad evaluation strategy, in combination with qualitative and participatory evaluations[33]. Another way to reduce the limitations arising from indicator use is to continually test and re-assess the indicators[34]. In the case of the CBR Indicators, a priority should be to do this in partnership with communities and people with disabilities in order to promote their uptake.

The use of PSM as a method for analysis of cross-sectional data collected from the CBR Indicators is conceptually strong, due to its ability to reduce bias due to confounding variables in observational studies[9]. However, the methodological limitations of PSM also need to be considered. PSM requires that each participant has a non-zero probability of receiving treatment, meaning only people with disabilities can be included in the analysis. Due to this, one of the main advantages of

the CBR Indicators, namely the ability to use comparison individuals from the community, is lost[4]. Furthermore, PSM only controls for known covariates which means that there is a potential for bias if some covariates that affect the outcome are not included[9]. For example, in this study no data were available on the ethnicity of participants, despite its known association with social disparities in Vietnam[35]. Another such covariate in this study could be disability severity, although this was partially adjusted for in both the participant selection, whereby all people with disabilities were identified using the same government disability criteria, and further in the analysis through the inclusion of the self-rated health covariate. Another limitation of PSM is that it leads to reduced sample size which could limit generalizability, though this is partly addressed through the provided sensitivity analysis. The reduced sample size also increases the risk of type II error[36], but the sample size of this study met the commonly recommended minimum sample size of $10(p + 1)$, where p is the number of matching variables[37]. This study presents a starting point to encourage the generation of quantitative CBR research and demonstrates one possible method for reducing bias when analyzing cross-sectional CBR data. Further studies should look into additional statistical methods for analyzing the results obtained from the CBR Indicators.

Based on the present study, we recommend the further use and testing of the WHO CBR Indicators to increase standardized data collection in the field of CBR. In accompaniment to increased data collection, we recommend PSM as a method to reduce bias in cross-sectional CBR data analyses, especially for international comparisons where differences between populations may be greater than the within country differences observed in this study. Since using cross-sectional data presents limitations even after adjusting for bias, we also emphasize the need for future longitudinal data collection in order to assess effectiveness in the field of CBR.

CONCLUSION

This study presents the first use of PSM as a method for analyzing cross-sectional CBR data. While randomized and longitudinal data are ideal for evaluations, cross-sectional data presents the advantage of being more feasible to collect and thereby providing an essential foundation to generate hypotheses and perform further studies. Therefore, it is essential that appropriate statistical methods are applied to capitalize on available data. The potential of using PSM for analyzing cross-sectional CBR data was demonstrated, though further research should investigate alternative inferential methods, such as cluster matching or adjusted regression, which may be more suitable in allowing for the comparison of the differences between persons with and without disabilities in line with the WHO CBR Indicators. We recommend that the questions and indicators be continually reviewed, and that future cross-sectional CBR studies use PSM to reduce bias when comparing groups.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING INTERESTS

The authors declare no competing interests.

DATA SHARING

The data are owned by the World Health Organization (WHO). Data are available from the WHO for researchers who meet the criteria for access to confidential data. Interested researchers can access the data by the same means the authors accessed them, by contacting WHO under disability@who.int. Statistical code is available from the corresponding author.

ACKNOWLEGEMENTS

The authors express gratitude to those from the Vietnamese Ministries and the local healthcare workers who lent their knowledge and field experience. Special thanks go out to the interviewers who took the time to visit and speak with the participants; in Huế: Nguyen Thi Phung Diem, Nguyen Van Hong, Thuong Thi Huong Giang, Thian Cong Chirh, and Nguyen Thi Ngoc Anh and in Hoa Binh: Ha Thi Thoan, Vu Dury Hieu, Le Tleaal Hoa, Nguyen Quoc Dung, Le Vai Huy, and Nguyen Thanh.

AUTHOR CONTRIBUTIONS

Conceptualization: CM JW CS
Investigation: CM JW
Methodology: CM JW CS
Data curation: CM DMT
Formal analysis: CM
Project administration: CM JW CS
Resources: CM JW CS DMT
Supervision: CM DMT CS JW
Writing – original draft: CM CS JW
Writing – review & editing: CM JW CS DMT

REFERENCES

1 International Labour Organization, United Nations Educational, Scientific and Cultural Organization and the World Health Organization. CBR: A strategy for rehabilitation, equalization of opportunities, poverty reduction and social inclusion of people with disabilities: joint position paper. Geneva: World Health Organization, 2004.
2 World Health Organization, United Nations Educational, Scientific and Cultural Organization, & International Labour Organization. Community-based rehabilitation: CBR Guidelines. Geneva: World Health Organization, 2010.
3 World Health Organization, International Disability and Development Consortium. Capturing the difference we make- Community-based rehabilitation indicators manual. Geneva: World Health Organization, 2015.
4 Mason C, Weber J, Atasoy S, et al. Development of indicators for monitoring Community-Based Rehabilitation. *PLoS One* 2017;12(6):e0178418.
5 Finkenflügel H, Wolffers I, Huijsman R. The evidence base for community-based rehabilitation: a literature review. *Int J Rehabil Res* 2005;28(3):187-201.
6 Cornielje H, Velema JP, Finkenflugel H. Community based rehabilitation programmes: Monitoring and evaluation in order to measure results. *Lepr Rev* 2008;79(1):36-49.
7 Iemmi V, Blanchet K, Gibson LJ, et al. Community-based rehabilitation for people with physical and mental disabilities in low- and middle-income countries: a systematic review and meta-analysis. *Campbell Sys Rev* 2015;11(15).

8 Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health* 2012:jech-2011-200375.
9 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55.
10 Heckman JJ, Ichimura H, Todd P. Matching as an econometric evaluation estimator. *Rev Econ Stud* 1998;65(2):261-94.
11 Rosenbaum PR. The role of known effects in observational studies. *Biometrics* 1989;45(2):557-69.
12 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Research* 2011;46(3):399-424.
13 Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J* 2012;33(15):1893-901.
14 Zhang Z, Ni H, Xu X. Observational studies using propensity score analysis underestimated the effect sizes in critical care medicine. *J Clin Epidemiol* 2014;67(8):932-39.
15 Li M. Using the Propensity Score Method to Estimate Causal Effects A Review and Practical Guide. *Organ Res Methods* 2013;16(2):188-226.
16 Weber J, Polack S, Hartley S. An Online Survey on Identification of Evaluation Capacity, Needs and Current Practice of Programme Evaluation in Community-based Rehabilitation. *Disability CBR Inclusive Development* 2016;27(2):5-18.
17 Grandisson M, Hébert M, Thibeault R. A systematic review on how to conduct evaluations in community-based rehabilitation. *Disabil Rehabil* 2014;36(4):265-75.
18 Becerril J, Abdulai A. The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach. *World Devel* 2010;38(7):1024-35.
19 Jalan J, Ravallion M. Estimating the benefit incidence of an antipoverty program by propensity-score matching. *J Bus Econ Stat* 2003;21(1):19-30.
20 Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci* 2010;25(1):1.
21 Wu S, Wang R, Zhao Y, et al. The relationship between self-rated health and objective health status: a population-based study. *BMC Public Health* 2013;13(1):320.
22 Pollack CE, Chideya S, Cubbin C, et al. Should health studies measure wealth?: a systematic review. *Am J Prev Med* 2007;33(3):250-64.
23 Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc* 2007;99(9):1013.
24 Garrido M, Kelley A, Parisc J, Roza K, Meier D, Morrison R, Aldridge M. (2014). Methods for constructing and assessing propensity scores. *Health Serv Res* 2014;49(5):1701-1720.
25 Stef van Buuren, Karin Groothuis-Oudshoorn). mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; 45(3):1-67. URL http://www.jstatsoft.org/v45/i03/.
26 Cochran WG, Rubin DB. Controlling bias in observational studies: A review. *Sankhya Ser A* 1973;417-46.
27 Ho DE, Imai K, King G, et al. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal* 2007;15(3):199-236.
28 Austin, Peter C. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med* 2008;27.12: 2037-2049.
29 Keele L. An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. R package version 2.1. https://CRAN.R-project.org/package=rbounds r, 2010.
30 Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011. 42(8):1-28. URL http://www.jstatsoft.org/v42/i08/).
31 Mauro V, Biggeri M, Deepak S, et al. The effectiveness of community-based rehabilitation programmes: an impact evaluation of a quasi-randomised trial. *J Epidemiol Community Health* 2014;68(11):1102-08.

32 Biggeri M, Deepak S, Mauro V, et al. Do community-based rehabilitation programmes promote the participation of persons with disabilities? A case control study from Mandya District, in India. *Disabil Rehabil* 2014;36(18):1508-17.

33 Weber J, Grech S, Polack S. Towards a 'mind map' for evaluative thinking in Community Based Rehabilitation: reflections and learning. *Disability Global South* 2016;3(2):951-79.

34 Implementing the vision: Addressing challenges to results-focused management and budgeting. Implementation Challenges in Results Focused Management and Budgeting; 2002 11-12 February; Paris. OECD.

35 Taylor P. Social inequality in Vietnam and the challenges to reform. Singapore: Institute of Southeast Asian Studies 2004:23,212.

36 Streiner, DL, Norman, GR. The pros and cons of propensity scores. *Chest* 2012;142(6):1380-1382.

37 Tumlinson, SE, Sass, DA, Cano, SM. The search for causal inferences: Using propensity scores post hoc to reduce estimation error with nonexperimental research. *J Pediatr Psychol* 2014;39(2):246-257.

Supplementary Table. The WHO CBR Survey questions, response options and analysis categories for the matching variables.

Variable	Survey Question and Response Options	Analysis Categories
Gender	Record the gender of the selected participant	Male=1, Female=0
Age	How old are you?	0-5yrs=1; 6-12yrs=2; 13-17yrs=3; 18-24yrs=4; 25-44yrs=5; 45-64yrs=6; 65+yrs=7
Education Level (for creation of SES)	What is the highest level of education you have achieved, or are working to achieve? *1=No schooling or never completed any grade; 2=Elementary education; 3=Vocational education; 4=Professional training; 5=Secondary school; 6=College; 7=University; 8=Post-graduate studies; 9=Other*	If respondent answered 1: score=0 If respondent answered 2: score=1 If respondent answered 3,4,5: score=2* If respondent answered 6,7,8: score=3
Employment Grade	What is your current working situation? *1=Not working and looking for work; 2=Not working for wages and not looking for paid work; 3=Working for wages or salary with an employer; 4=Working for wages, but currently on sick leave; 5=Self-employed or own-account worker; 6=Working as unpaid family member; 7=Retired because of the health condition; 8=Retired because of age; 9=Early retirement; 10=Other*	If respondent answered 1: score=0 If respondent answered 6: score=1 If respondent answered 2,3,4,5,7,8,9: score=2 If respondent answered 10: score=missing
Health Status (for creation of SES)	In general, how would you rate your health today? *1 = Very good; 2 = Good; 3 = Neither poor nor good; 4 = Poor; 5 = Very poor*	Inverted
Received needed medical care	In the last 12 months, has there been a time when you needed health care but did not get that care? *1 = Yes; 2 = No; 3 = No need for health care in the past 12 months*	If respondent answered 1: score=0 If respondent answered 2 or 3: score=1
Involved in making treatment decisions	On your last visit to a health care provider, to what extent were you involved in making decisions for your treatment? *1 (Not at all); 2; 3; 4; 5 (Completely)*	No transformation
Received needed rehabilitation services	In the last 12 months, has there been a time when you needed rehabilitation services, such as physical, occupational, or speech therapy, but did not get those services? *1 = Yes; 2 = No; 3 = No need for rehabilitation services in the past 12 months*	If respondent answered 1: score=0 If respondent answered 2 or 3: score=1
Aware of financial services	Do you know how to get financial services such as credit, insurance, grants, and savings programs? *1 = Yes; 2 = No*	If respondent answered 1: score=1 If respondent answered 2: score=0
Receive social protection	Do you currently benefit from any social protection program, such as loss of income through old age, sickness or disability? *1 = Yes; 2 = No*	If respondent answered 1: score=1 If respondent answered 2: score=0
Participation in self-help group	Are you a member of a self-help group? *1=Yes; 2=No, but I would like to; 3=No, I don't want to*	If respondent answered 1 or 3: score=1 If respondent answered 2: score=0

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract Pg 1, 1-3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found Pg 2, 3-28
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported Pg 4, 2-39
Objectives	3	State specific objectives, including any prespecified hypotheses Pg 4, 40- pg 5, 3
Methods		
Study design	4	Present key elements of study design early in the paper Pg 5, 7-8; 14
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection Pg 5, 14-37
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants Pg 5, 27; pg 8, 5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable Pg 5, 40- pg 6, 23; Supplementary Table
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group Pg 5, 7-13
Bias	9	Describe any efforts to address potential sources of bias Pg 7, 20-27 → The analysis uses PSM, a method to reduce bias
Study size	10	Explain how the study size was arrived at Pg 8, 2-5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why Pg 5, 40- pg 6, 23; Supplementary Table
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding

	Pg 6, 6-29
(*b*) Describe any methods used to examine subgroups and interactions	
(*c*) Explain how missing data were addressed Pg 7, 1-2	
(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
(*e*) Describe any sensitivity analyses Pg 7, 23-27; pg 9, 8-14	

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed Pg 8, 2-4
		(b) Give reasons for non-participation at each stage Pg 8, 7-8
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders Pg 8, 9-10, 21-25
		(b) Indicate number of participants with missing data for each variable of interest Pg 7, 1-2
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures Pg 5, 40- pg 6, 2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included Pg 9, 1-5
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses Pg 9, 5-14
Discussion		
Key results	18	Summarise key results with reference to study objectives Pg 9, 20-26
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias Pg 10, 7-22; 29-33; pg 10, 38- pg 11, 1-7
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence Pg 11, 8-18
Generalisability	21	Discuss the generalisability (external validity) of the study results Pg 11, 4-6
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based Pg 11, 32-33

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Can propensity score matching be applied to cross-sectional data to evaluate Community-based Rehabilitation? Results of a survey implementing the World Health Organization's Community-based Rehabilitation Indicators in Vietnam

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-022544.R2
Article Type:	Research
Date Submitted by the Author:	07-Nov-2018
Complete List of Authors:	Mason, Catherine; LMU Munich, Department for Medical Information Processing, Biometry and Epidemiology (IBE) Sabariego, Carla; Ludwig-Maximilians University, Public Health and Health Services Research Thắng, Đoàn Mạnh ; Hoa Binh Department of Health Weber, Jörg; CBM eV
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Global health, Public health, Epidemiology
Keywords:	epidemiological methods, disability, social inequalities

Can propensity score matching be applied to cross-sectional data to evaluate Community-based Rehabilitation? Results of a survey implementing the World Health Organization's Community-based Rehabilitation Indicators in Vietnam

Authors: Catherine Mason[1*], Carla Sabariego[1], Đoàn Mạnh Thắng[2], Joerg Weber[3]

[1] Department for Medical Information Processing, Biometry and Epidemiology (IBE)
Ludwig-Maximilians-Universität München
Marchioninistraße 15
81377 Munich, Germany

Catherine.Mason@med.uni-muenchen.de
+49 (0)89 2180-78229

[2] Hoa Binh Department of Health (Sở Y tế tỉnh Hòa Bình)
Hoa Binh City, Vietnam

[3] CBM International
Bensheim, Germany

Keywords: disability, epidemiological methods, social inequalities

Word count: 3787

ABSTRACT

Objectives: Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalize opportunities and include people with disabilities in all aspects of life. The complexity of CBR and often limited resources lead to challenges when attempting to quantify its effectiveness, with randomization and longitudinal data rarely possible. Statistical methods, such as propensity score matching (PSM), offer an alternative approach to evaluate a treatment when randomization is not feasible. The aim of this study is to examine whether PSM can be an effective method to facilitate evaluations of results in CBR when data are cross-sectional.

Design: Cross-sectional survey

Setting and Participants: Data were collected using the World Health Organization's CBR Indicators in Vietnam, with treatment assignment (participating in CBR or not) determined by province of residence. 298 participants were selected through government records.

Results: PSM was conducted using one-to-one nearest neighbour method on ten covariates. In the unmatched sample, significant differences between groups were found for six of the ten covariates. PSM successfully adjusted for bias in all covariates in the matched sample (74 matched pairs). A paired t-test compared the outcome of "community inclusion" (a score based on selected indicators) between CBR and non-CBR participants for both the matched and unmatched samples, with CBR participants found to have significantly worse community inclusion scores (mean=17.86, SD=6.30, 95%CI 16.45-19.32) than non-CBR participants (mean=20.93, SD=6.16, 95%CI 19.50-22.35); t(73)=3.068, p=0.001. This result did not differ between the matched and unmatched samples.

Conclusion: PSM successfully reduced bias between groups, though its application did not affect the tested outcome. PSM should be considered when analyzing cross-sectional CBR data, especially for international comparisons where differences between populations may be greater.

Strengths and limitations of this study

- The complexity of CBR and often limited resources available in the field lead to challenges in research attempting to quantify its effectiveness and to a heavy reliance on non-randomized cross-sectional data, implying the need for statistical approaches, such as PSM, to account for these limitations.
- PSM attempts to mimic randomization by creating a sample of participants who received the treatment (CBR participants) that is comparable on all observed covariates to participants who did not receive the treatment (non-CBR participants).
- The potential of using PSM for analyzing cross-sectional CBR data was demonstrated, as biases detected in the distribution of covariates between groups in the unmatched samplewere successfully eliminated.
- One of the main advantages of the CBR Indicators, namely the ability to use comparison individuals without disability from the community is lost; as PSM requires that all participants have a non-zero probability of receiving treatment meaning only people with disabilities can be included.

- PSM only controls for known covariates which means that there is a potential for bias if some covariates that affect the outcome are not included.

INTRODUCTION

Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalize opportunities and include people with disabilities in all aspects of community life. It is broadly defined as "a strategy within general community development for the rehabilitation, equalization of opportunities and social inclusion of all people with disabilities"[1]. The wide scope of CBR is further expanded through the various implementing stakeholders involved in CBR, including people with disabilities themselves, their families and communities, and the relevant governmental and non-governmental service sectors. It is due, at least in part, to this extensive definition that reliable and internationally comparable data to monitor and evaluate CBR are scarce. In an effort to synthesize global perspectives on CBR, the World Health Organization (WHO) developed their "Community-Based Rehabilitation Guidelines" in 2010 which have since become accepted as a conceptual framework for CBR[2]. With these guidelines, WHO emphasized the need for a common global framework for monitoring CBR in line with the Convention on the Rights of Persons with Disability (CRPD). With the launch of the global WHO CBR Indicators in 2015, there is now a standardized approach to do this[3,4].

The complexity of CBR leads to challenges in research attempting to quantify its effectiveness[5-7]. Fully experimental studies with randomization are rarely possible for both ethical and practical reasons, which inherently lead to limitations. The possibility of bias arises as the apparent difference in an outcome between two treatment groups may depend on characteristics that affected whether or not an individual received a given treatment, instead of being an actual effect of the treatment. For this reason there has been a recent emphasis on so-called natural experiments, where a range of primarily statistical approaches are used to evaluate a treatment or intervention when randomization is not feasible[8]. One such approach is propensity score matching (PSM).

PSM was first presented in 1983 by Rosenbaum and Rubin as a method to reduce bias due to confounding variables in observational studies[9]. It attempts to mimic randomization by creating a sample of participants who received the treatment that is comparable on all observed covariates to participants who did not receive the treatment. This effectively creates an experimental dataset where the comparison group is, on average, equivalent to individuals in the exposed group on all observed covariates[10-12]. A systematic review comparing 21 PSM studies to 63 RCTs on therapeutic interventions for acute coronary syndromes found that PSM produced more extreme treatment effect estimates when compared with those from RCTs, although these differences were rarely statistically significant[13]. A similar comparison including 20 propensity-score-based studies matched to RCT results was conducted examining critical care medicine and found that propensity-score-based studies report less beneficial effects of treatment in comparison to RCTs[14]. Despite some shortcomings, PSM provides a method for evaluating complex interventions where randomization is not possible.

PSM has been increasingly used in various research fields, including Public Health, to evaluate complex interventions[15]. CBR is considered a complex intervention, and data collection in the field is further hindered by low resources making quantitative longitudinal data collection infeasible and rarely done[6,7,16,17]. This implies that data analysis in the field of CBR relies heavily on cross-sectional data. PSM has already been successfully applied to cross-sectional data[18,19]. Therefore, the main objective of this paper is to examine whether PSM can be an effective method to facilitate evaluations of results in CBR when data are cross-sectional. Data used in the present study were collected using the WHO CBR Indicators in Vietnam in 2016 with the assignment of persons to the treatment (CBR participants) and non-treatment group (non-CBR participants) determined by province of residence.

PSM will be conducted on the outcome of *community inclusion* of people with disabilities, the ultimate goal of CBR in strong alinement with the CRPD, using a sum score of WHO CBR social indicators and an empowerment indicator.

METHODS

Data Collection

Data collection was conducted using the survey questionnaire accompanying the WHO CBR Indicators[3]. These indicators examine differences in health, education, social life, livelihood and empowerment between people with disabilities and other community members. There are two subsets of indicators: base indicators which are broad and should be used in all data collection activities to ensure comparability, and supplementary indicators which can provide more specific coverage, and can be selected depending on the specific CBR goals and strategies of a program. The indicators and corresponding questions used in this paper are presented in table 1.

This study presents a secondary analysis of data collected during a multi-site cross-sectional survey in 2016 in two Vietnamese provinces: Huế, where CBR is fully implemented and all districts have CBR coverage through government implementation and through non-governmental organizations' (NGO) activities; and Hòa Bình, where CBR is not implemented by either government or NGOs. The Huế CBR program began in 2009 in cooperation with the Huế Rehabilitation Hospital. The program focused mainly on activities to increase capacity building for CBR workers, not only in terms of rehabilitation skills, but also working to improve their counselling and networking skills. The other focus of the program was to strengthen referral pathways for people with disabilities so that they could be connected with other existing services in the province, such as schools with teachers who were trained to support students with disabilities and vocational training centers. An Android mobile phone application (app), available from WHO for the CBR Indicators, was used to collect data during interviews (app free to download at: http://play.google.com/store/apps/details?id=com.universaltools.whocbrsurvey&hl=en).

People with disabilities were identified prior to the survey by government records. In both provinces a team of five local health care workers were trained by the lead researcher (CM) over two days on how to conduct interviews using the survey questions and the app. Data collection was supervised by CM. Data were collected during face-to-face interviews with data recorded anonymously. All respondents were informed of the purpose of the study, and then provided verbal (Huế) or written consent (Hòa Bình). In Huế the decision to provide verbal rather than written consent was justified since requiring written consent would embarrass illiterate participants, leading to a decreased willingness to answer further questions truthfully. In instances when the respondent had cognitive limitations that prevented them from being interviewed, or if the respondent was a minor, a proxy interview with a family member was performed. Ethical approval was obtained through the Ludwig-Maximilians-Universität Munich Ethics Commission and by the local provincial Ministries of Health.

Variables

Outcome Variable

To measure community inclusion, a sum score was created from the social base and supplementary questions, with the addition of the base question from empowerment. These questions all used the

same response scale of *1(Not at all)* to *5(Completely)* with the final sum score ranging from 4 to 33, with higher scores indicating higher levels of inclusion (table 1).

Table 1. WHO CBR Indicators and questions used to measure them. Base indicators are shown in bold. The response option for all questions ranged from 1 (Not at all) to 5 (Completely).

Component	Indicator	Survey Question
Social	**% of people with disability that feel valued as individuals by members of their community**	Do you feel that other people respect you? For example, do you feel that others value you as a person and listen to what you have to say?
	% of people with disability who make their own decisions about the personal assistance they need	Do you get to make decisions about the personal assistance that you need (who assists you, what type of assistance, when to get assistance)?
	% of people with disability make their own decisions about their personal relationships	Do you get to make your own decisions about your personal relationships, such as friends and family?
	% of people with disability who participate in artistic, cultural or religious activities	Do you get to participate in artistic, cultural or religious activities?
	% of people with disability who participate in mainstream recreational, leisure and sports activities	Do you get to participate in community recreational, leisure and sports activities?
	% of people with disability who know their legal rights	To what extent do you know your legal rights?
Empower-ment	**% of people with disability who make informed choices and decisions**	Do you get to make the big decisions in your life? For example, deciding who to live with, where to live, or how to spend your money?

Matching Variables

Matching variables were those available from the WHO CBR Indicators, and were selected based on their theoretical association with community inclusion and CBR group assignment, primarily using CBR Guidelines[2]. Data on *age* and *gender* were collected. Age was collected in categories (see table 2) which were dichotomized for the analysis[20]. Though data on disability severity were not available, *general health status* was used as a proxy, using the question *"How would you rate your health today?"*[21]. A variable for *socio-economic status* (SES) was created using a sum score based on the questions "*What is the highest level of education you have achieved or are working to achieve?"* and "*Do you have enough money to meet your needs?".* The first question is commonly used in SES variable creation, and the second question targets wealth[22,23]. The variable *province of residence* corresponded to CBR coverage (no coverage in Hòa Bình, full CBR coverage in Huế). To account for economic differences between the provinces that might not be captured by SES, the covariate *receiving social protection* (such as for loss of income through old age, sickness or disability) was included. Covariates of *financial awareness* (knowing how to get financial services or social protection if needed), *having access to health services when needed,* and *having access to rehabilitation services when needed* were also included. A proxy for autonomy was captured through the covariates of *being involved in decision making regarding medical treatment* and *participating in a self-help group if desired* (see supplementary table). Seeing as the CBR program in Huế focused on increasing referral pathways within the medical and education sectors, the questions derived from the education component and many from the medical component were not included as matching variables, since including covariates associated with CBR participation but not with community inclusion decrease model precision[24].

Missing Data

Missing data were low (2.25%). Multiple imputation (five imputations) using fully conditional specification (MICE package [25] in R Studio Version 0.99.903) was used to replace missing data.

Analysis

Matching on the Propensity Score

The number of treated and untreated participants were similar (difference of n=4). Therefore, participants were matched using one-to-one nearest neighbour technique, which matched each treated unit to one control that was closest using calipers of width equal to 0.25 of the standard deviation (SD) of the logit of the estimated propensity score without iteration[26]. This implies that for a given treated participant, all the untreated participants are identified whose scores are within this specified distance and then the best match is formed. If no match falls within this distance the participant is excluded. Participants were matched on ten covariates (see *Matching Variables*).

Balance Diagnostics

Baseline comparisons between the covariates were conducted for the matched and unmatched samples. Balance diagnosis was performed using the standardized difference method, which compares the difference in means of each covariate in units of the pooled SD for the matched and unmatched samples[12]. Successful matching is indicated when the absolute standardized differences of means is less than 0.25[27].

Comparing Groups

For the community inclusion outcome, data matched on the ten covariates were compared using a paired t-test[28]. Bootstrapping was performed (1000 samples) in order to produce 95% confidence intervals (CI), which has been shown to account for uncertainty in the matching procedure[20].
A sensitivity analysis was performed using the Rosenbaum Bounds for Hodges-Lehmann Point Estimate to assess how robust the findings were to hidden bias due to unobserved covariates ('rbounds' package [29] in R Studio Version 0.99.903). The maximum Gamma (the odds of differential assignment to treatment due to unobserved factors) was set to 2 with increments of 0.1 to test at which point the between group differences are no longer robust[29].

Data cleaning was performed using SPSS version 23 (copyright IBM Corporation). PSM was performed in R Studio (Version 0.99.903) using the 'MatchIt' package[30].

Patient and Public Involvement

Participants were not directly involved in the development of the research question, study design, recruitment or conduct of the study. However, in the province of Huế (where CBR is implemented) participants are continually involved in the development of the CBR program, as CBR is participatory in nature. It was through their motivation – stemming from the need to prove to the national government and international donors that their intervention has an impact in order to receive funds – that the survey was conducted in the first place. A study report was submitted to the Huế and Hòa Bình Ministries of Health which presented simple numeric and graphic descriptive findings which were to be communicated to participants.

RESULTS

Data were available from 298 participants. In Huế, 575 people with disabilities were identified by government records and 147 were included, while in Hòa Bình 375 people were identified by government records and 151 were included (sample size calculated using an alpha significance level of 0.05 and power of 90%). Included participants were randomly selected from the complete list. After the random selection, each interviewer was assigned a group of selected participants based on their geographic location. Of the randomly selected participants, one in Hòa Bình could not be contacted so another participant was selected. In both provinces none of the invited participants refused participation. Males comprised 153(51.3%) of the participants, with a modal age group of 45-64(28.9%) (see table 2 for further descriptives).

In the unmatched sample, CBR participants had higher health status, were more likely to participate in a self-help group, more financially aware and more likely to be receiving social protection while they had worse access to rehabilitation services. Some age differences were also noted (table 2). In the unmatched sample the absolute standardized difference across the 10 covariates ranged from 0.008 to 1.008 indicating bias.

When CBR participants were matched with non-CBR participants on the logit of the specified propensity score model, 74 matched pairs were formed. This meant that 49.7% of CBR participants were successfully matched to a control. PSM was successful in reducing bias between the covariates in the matched sample, as the standardized differences ranged from 0 to 0.147 with all values falling below the threshold value of 0.25[27] (table 2).

Table 2. Baseline characteristics of CBR participants and non-CBR participants in the unmatched and matched samples. Absolute standardized differences of means are shown, with differences exceeding the threshold of 0.25 indicated in bold.

Variable		Unmatched Sample			Matched Sample		
		No CBR (n=151)	With CBR (n=147)	Std. dif. of means	No CBR (n=74)	With CBR (n=74)	Std. dif. of means
Age	0-5	11 (7.2%)	6 (4.1%)	0.161	3 (4.1%)	5 (6.8%)	0.136
	6-12	19 (12.6%)	11 (7.5%)	0.193	7 (9.5%)	5 (6.8%)	0.102
	13-17	4 (2.6%)	6 (4.1%)	0.072	2 (2.7%)	1 (1.4%)	0.068
	18-24	12 (7.9%)	12 (8.2%)	0.008	7 (9.5%)	7 (9.5%)	0.000
	25-44	49 (32.5%)	32 (21.8%)	**0.258**	23 (31.1%)	22 (29.7%)	0.033
	45-64	42 (27.8%)	44 (29.9%)	0.046	21 (28.4%)	26 (35.1%)	0.147
	65+	14 (9.3%)	36 (24.5%)	**0.353**	11 (14.9%)	8 (10.8%)	0.094
Gender (male)		80 (53.0%)	73 (50.0%)	0.066	37 (50.0%)	42 (56.8%)	0.135
SES (range 1-10)		3.74±1.32	3.91±1.30	0.235	3.65±1.45	3.67±1.42	0.020
Health status (range 1-5)		2.89±0.77	3.37±0.70	**0.683**	3.05±0.75	3.14±0. 65	0.115
Receiving social protection		74 (49.0%)	117 (79.6%)	**1.008**	48 (64.9%)	52 (70.3%)	0.141
Access to health services		132 (87.4%)	126 (85.7%)	0.048	66 (89.2%)	66 (89.2%)	0.000
Access to rehabilitation services		128 (84.8%)	123 (83.7%)	**0.263**	29 (39.2%)	31 (41.9%)	0.054
Self-help group		63 (41.7%)	75 (51.0%)	**0.396**	31 (41.9%)	32 (43.2%)	0.027
Financial awareness		73 (48.3%)	122 (83.0%)	**0.789**	51 (68.9%)	55 (74.3%)	0.134
Involved in treatment decisions		47 (31.1%)	65 (44.2%)	0.137	65 (87.8%)	65 (87.8%)	0.000

Note: continuous variables are presented as means ± standard deviation; dichotomous variables are presented as n(%)

To test whether PSM affected the pre-defined outcome of community inclusion, the difference between groups in the matched and unmatched samples were assessed: similar significant differences were found. In the matched sample, CBR participants had worse community inclusion scores

(mean=17.86, SD=6.30, 95%CI 16.45-19.32) than non-CBR participants (mean=20.93, SD=6.16, 95%CI 19.50-22.35); t(73)=3.068, p=0.001. The sensitivity analysis corroborated the results, showing that CBR participants had a median difference in community inclusion score 3.5 points lower than non-CBR participants (Gamma=0). When the Gamma value was increased to 2, the upper and lower bounds did not include zero indicating robust results[29]. In a further sensitivity analysis, to ensure that the covariate of "access to rehabilitation" did not bias the model by being more strongly associated with receiving CBR rather than with the outcome of community inclusion, the model was run excluding this variable. The new model resulted in 75 matched pairs with all standardized differences falling below the threshold. The results of the t-test did not differ from the model including access to rehabilitation; CBR participants had worse community inclusion scores (mean=18.11, SD=5.981, 95%CI 16.72-19.47) than non-CBR participants (mean=21.17, SD=6.381, 95%CI 19.67-22.60); t(74)=3.310, p=0.0014.

Overall, the results did not differ from the results before PSM: community inclusion for participants with CBR (mean=18.61, SD=5.38) and without CBR (mean=20.64, SD=6.49); t(296)=2.935, p=0.004 using an independent t-test.

DISCUSSION

To our knowledge, this study presents the first use of PSM as a method for analyzing cross-sectional data in the field of CBR. The study analyzed data collected using the WHO CBR Indicators, and found that community inclusion scores of CBR participants were significantly lower than those of non-CBR participants after PSM. Despite bias being detected in the distribution of covariates between groups in the unmatched sample, the results before PSM did not significantly differ from those after. We conclude that PSM can be successfully applied to cross-sectional CBR data, though in this case the bias reduction provided by PSM did not affect the tested outcome.

PSM has been applied only to longitudinal CBR data so far, but PSM studies using cross-sectional data are available from other fields. These studies had similar results in terms of the methodological success of PSM, but unlike our study they had final outcomes in line with their hypotheses. One such example is the study from Jalan and Ravallion which examines the effect of an employment-based poverty reduction program on income gain, accounting for pre-intervention and foregone income[19]. Through the trial of three PSM methods, they were able to reduce the differences between the two populations and to demonstrate the effectiveness of the program. Another such example is the study from Becerril and Abdulai showing the positive impact of new maize farming technologies on per capita poverty outcomes[18]. Similarly to our study, they detected bias in the distribution of covariates between groups in the unmatched sample, indicating that accounting for bias though PSM was important. In the field of CBR, PSM has been used to evaluate longitudinal CBR data in India, looking at livelihood and health outcomes[31,32]. PSM was used to reduce the bias between the CBR and non-CBR groups, with results showing that CBR participants had better health and livelihood outcomes, and that these differences generally increased over time at both four years and seven years. In our study, data was collected seven years after the program began, which would make the timing comparable and it is therefore plausible that the effect of CBR in our study could already be quantifiable. As in our study, these studies all showed bias between unmatched groups which were reduced in the matched sample after PSM. However, none of these studies presented their outcome results of the unmatched sample for comparison, so it cannot be determined if their final results were unaffected by matching as is the case in our study.

The results of the present study go against the anecdotal evidence that CBR has a positive influence on the lives of people with disabilities[6,7,33]. Results from longitudinal data indicate that CBR has a positive impact on receiving pensions, accessing paid jobs, accessing assistive devices, and personal-practical autonomy, with the impact increasing over time[31]. An explanation for our results could be that cross-sectional data allow for comparisons between groups at a single time-point, and even after PSM is applied to reduce bias the causal relationship between CBR implementation and social inclusion cannot be determined. While the cross-sectional data collected in this study represent the first quantitative data from the region and therefore an important foundation for future work, the results emphasizes the general need for further collection and publication of CBR data, especially longitudinal data. Additionally, this study focused on community inclusion - the ultimate goal of CBR - but when interpreting results it is also important to consider the specific targets of the program being examined. Though CBR aims to impact all aspects of the lives of people with disabilities to increase community inclusion, the program in Huế does not directly target community inclusion. The program focuses on increasing the capacity of CBR workers and on strengthening referral pathways with the medical and educational sectors. Through these activities the community inclusion of people with disabilities should improve over time, but since community inclusion was not the direct target of the program, the community inclusion effects might only appear after a longer period, which could be a reason for the counter-intuitive results. Therefore, when assessing a program in its early stages, it may be more important to match the indicators used with the specific targets of programs.

To our knowledge, this study is the first to implement the recently developed WHO CRB Indicators[4]. The study highlights how important it is to collect standardized data in the field of CBR in order to facilitate comparisons between groups and determine effectiveness of programs. One of the main advantages of the CBR Indicators and their data collection strategy is that they are easy to use in the field. The indicators allow for descriptive comparisons to be made easily, but for indicators to be used appropriately it is important to go beyond these descriptive results using inferential statistics. Furthermore, no single indicator or even a set of indicators is capable of capturing all changes in dynamic settings. The use of indicators alone has the potential limitation of collecting meaningless or misleading information,[34] and therefore they should be used as part of a broad evaluation strategy, in combination with qualitative and participatory evaluations[33]. Another way to reduce the limitations arising from indicator use is to continually test and re-assess the indicators[34]. In the case of the CBR Indicators, a priority should be to do this in partnership with communities and people with disabilities in order to promote their uptake.

The use of PSM as a method for analysis of cross-sectional data collected from the CBR Indicators is conceptually strong, due to its ability to reduce bias due to confounding variables in observational studies[9]. However, the methodological limitations of PSM also need to be considered. PSM requires that each participant has a non-zero probability of receiving treatment, meaning only people with disabilities can be included in the analysis. Due to this, one of the main advantages of the CBR Indicators, namely the ability to use comparison individuals from the community, is lost[4]. Furthermore, PSM only controls for known covariates which means that there is a potential for bias if some covariates that affect the outcome are not included[9]. For example, in this study no data were available on the ethnicity of participants, despite its known association with social disparities in Vietnam[35]. Another such covariate in this study could be disability severity, although this was partially adjusted for in both the participant selection, whereby all people with disabilities were identified using the same government disability criteria, and further in the analysis through the

inclusion of the self-rated health covariate. Another limitation of PSM is that it leads to reduced sample size which could limit generalizability, though this is partly addressed through the provided sensitivity analysis. The reduced sample size also increases the risk of type II error[36], but the sample size of this study met the commonly recommended minimum sample size of 10(p + 1), where p is the number of matching variables[37]. This study presents a starting point to encourage the generation of quantitative CBR research and demonstrates one possible method for reducing bias when analyzing cross-sectional CBR data. Further studies should look into additional statistical methods for analyzing the results obtained from the CBR Indicators.

Based on the present study, we recommend the further use and testing of the WHO CBR Indicators to increase standardized data collection in the field of CBR. In accompaniment to increased data collection, we recommend PSM as a method to reduce bias in cross-sectional CBR data analyses, especially for international comparisons where differences between populations may be greater than the within country differences observed in this study. Since using cross-sectional data presents limitations even after adjusting for bias, we also emphasize the need for future longitudinal data collection in order to assess effectiveness in the field of CBR.

CONCLUSION

This study presents the first use of PSM as a method for analyzing cross-sectional CBR data. While randomized and longitudinal data are ideal for evaluations, cross-sectional data presents the advantage of being more feasible to collect and thereby providing an essential foundation to generate hypotheses and perform further studies. Therefore, it is essential that appropriate statistical methods are applied to capitalize on available data. The potential of using PSM for analyzing cross-sectional CBR data was demonstrated, though further research should investigate alternative inferential methods, such as cluster matching or adjusted regression, which may be more suitable in allowing for the comparison of the differences between persons with and without disabilities in line with the WHO CBR Indicators. We recommend that the questions and indicators be continually reviewed, and that future cross-sectional CBR studies use PSM to reduce bias when comparing groups.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING INTERESTS

The authors declare no competing interests.

DATA SHARING

The data are owned by the World Health Organization (WHO). Data are available from the WHO for researchers who meet the criteria for access to confidential data. Interested researchers can access the data by the same means the authors accessed them, by contacting WHO under disability@who.int. Statistical code is available from the corresponding author.

ACKNOWLEGEMENTS

The authors express gratitude to those from the Vietnamese Ministries and the local healthcare workers who lent their knowledge and field experience. Special thanks go out to the interviewers who took the time to visit and speak with the participants; in Huế: Nguyen Thi Phung Diem, Nguyen Van Hong, Thuong Thi Huong Giang, Thian Cong Chirh, and Nguyen Thi Ngoc Anh and in Hoa Binh: Ha Thi Thoan, Vu Dury Hieu, Le Tleaal Hoa, Nguyen Quoc Dung, Le Vai Huy, and Nguyen Thanh.

AUTHOR CONTRIBUTIONS

Conceptualization: CM JW CS
Investigation: CM JW
Methodology: CM JW CS
Data curation: CM DMT
Formal analysis: CM
Project administration: CM JW CS
Resources: CM JW CS DMT
Supervision: CM DMT CS JW
Writing – original draft: CM CS JW
Writing – review & editing: CM JW CS DMT

REFERENCES

1 International Labour Organization, United Nations Educational, Scientific and Cultural Organization and the World Health Organization. CBR: A strategy for rehabilitation, equalization of opportunities, poverty reduction and social inclusion of people with disabilities: joint position paper. Geneva: World Health Organization, 2004.
2 World Health Organization, United Nations Educational, Scientific and Cultural Organization, & International Labour Organization. Community-based rehabilitation: CBR Guidelines. Geneva: World Health Organization, 2010.
3 World Health Organization, International Disability and Development Consortium. Capturing the difference we make- Community-based rehabilitation indicators manual. Geneva: World Health Organization, 2015.
4 Mason C, Weber J, Atasoy S, et al. Development of indicators for monitoring Community-Based Rehabilitation. *PLoS One* 2017;12(6):e0178418.
5 Finkenflügel H, Wolffers I, Huijsman R. The evidence base for community-based rehabilitation: a literature review. *Int J Rehabil Res* 2005;28(3):187-201.
6 Cornielje H, Velema JP, Finkenflugel H. Community based rehabilitation programmes: Monitoring and evaluation in order to measure results. *Lepr Rev* 2008;79(1):36-49.
7 Iemmi V, Blanchet K, Gibson LJ, et al. Community-based rehabilitation for people with physical and mental disabilities in low- and middle-income countries: a systematic review and meta-analysis. *Campbell Sys Rev* 2015;11(15).
8 Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health* 2012:jech-2011-200375.
9 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55.
10 Heckman JJ, Ichimura H, Todd P. Matching as an econometric evaluation estimator. *Rev Econ Stud* 1998;65(2):261-94.
11 Rosenbaum PR. The role of known effects in observational studies. *Biometrics* 1989;45(2):557-69.
12 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Research* 2011;46(3):399-424.

13 Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J* 2012;33(15):1893-901.
14 Zhang Z, Ni H, Xu X. Observational studies using propensity score analysis underestimated the effect sizes in critical care medicine. *J Clin Epidemiol* 2014;67(8):932-39.
15 Li M. Using the Propensity Score Method to Estimate Causal Effects A Review and Practical Guide. *Organ Res Methods* 2013;16(2):188-226.
16 Weber J, Polack S, Hartley S. An Online Survey on Identification of Evaluation Capacity, Needs and Current Practice of Programme Evaluation in Community-based Rehabilitation. *Disability CBR Inclusive Development* 2016;27(2):5-18.
17 Grandisson M, Hébert M, Thibeault R. A systematic review on how to conduct evaluations in community-based rehabilitation. *Disabil Rehabil* 2014;36(4):265-75.
18 Becerril J, Abdulai A. The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach. *World Devel* 2010;38(7):1024-35.
19 Jalan J, Ravallion M. Estimating the benefit incidence of an antipoverty program by propensity-score matching. *J Bus Econ Stat* 2003;21(1):19-30.
20 Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci* 2010;25(1):1.
21 Wu S, Wang R, Zhao Y, et al. The relationship between self-rated health and objective health status: a population-based study. *BMC Public Health* 2013;13(1):320.
22 Pollack CE, Chideya S, Cubbin C, et al. Should health studies measure wealth?: a systematic review. *Am J Prev Med* 2007;33(3):250-64.
23 Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc* 2007;99(9):1013.
24 Garrido M, Kelley A, Parisc J, Roza K, Meier D, Morrison R, Aldridge M. (2014). Methods for constructing and assessing propensity scores. *Health Serv Res* 2014;49(5):1701-1720.
25 Stef van Buuren, Karin Groothuis-Oudshoorn). mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; 45(3):1-67. URL http://www.jstatsoft.org/v45/i03/.
26 Cochran WG, Rubin DB. Controlling bias in observational studies: A review. *Sankhya Ser A* 1973;417-46.
27 Ho DE, Imai K, King G, et al. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal* 2007;15(3):199-236.
28 Austin, Peter C. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med* 2008;27.12: 2037-2049.
29 Keele L. An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. R package version 2.1. https://CRAN.R-project.org/package=rbounds r, 2010.
30 Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011. 42(8):1-28. URL http://www.jstatsoft.org/v42/i08/).
31 Mauro V, Biggeri M, Deepak S, et al. The effectiveness of community-based rehabilitation programmes: an impact evaluation of a quasi-randomised trial. *J Epidemiol Community Health* 2014;68(11):1102-08.
32 Biggeri M, Deepak S, Mauro V, et al. Do community-based rehabilitation programmes promote the participation of persons with disabilities? A case control study from Mandya District, in India. *Disabil Rehabil* 2014;36(18):1508-17.
33 Weber J, Grech S, Polack S. Towards a 'mind map' for evaluative thinking in Community Based Rehabilitation: reflections and learning. *Disability Global South* 2016;3(2):951-79.
34 Implementing the vision: Addressing challenges to results-focused management and budgeting. Implementation Challenges in Results Focused Management and Budgeting; 2002 11-12 February; Paris. OECD.
35 Taylor P. Social inequality in Vietnam and the challenges to reform. Singapore: Institute of Southeast Asian Studies 2004:23,212.
36 Streiner, DL, Norman, GR. The pros and cons of propensity scores. *Chest* 2012;142(6):1380-1382.

37 Tumlinson, SE, Sass, DA, Cano, SM. The search for causal inferences: Using propensity scores post hoc to reduce estimation error with nonexperimental research. *J Pediatr Psychol* 2014;39(2):246-257.

Supplementary Table. The WHO CBR Survey questions, response options and analysis categories for the matching variables.

Variable	Survey Question and Response Options	Analysis Categories
Gender	Record the gender of the selected participant	Male=1, Female=0
Age	How old are you?	0-5yrs=1; 6-12yrs=2; 13-17yrs=3; 18-24yrs=4; 25-44yrs=5; 45-64yrs=6; 65+yrs=7
Education Level (for creation of SES)	What is the highest level of education you have achieved, or are working to achieve? *1=No schooling or never completed any grade; 2=Elementary education; 3=Vocational education; 4=Professional training; 5=Secondary school; 6=College; 7=University; 8=Post-graduate studies; 9=Other*	If respondent answered 1: score=0 If respondent answered 2: score=1 If respondent answered 3,4,5: score=2* If respondent answered 6,7,8: score=3
Employment Grade	What is your current working situation? *1=Not working and looking for work; 2=Not working for wages and not looking for paid work; 3=Working for wages or salary with an employer; 4=Working for wages, but currently on sick leave; 5=Self-employed or own-account worker; 6=Working as unpaid family member; 7=Retired because of the health condition; 8=Retired because of age; 9=Early retirement; 10=Other*	If respondent answered 1: score=0 If respondent answered 6: score=1 If respondent answered 2,3,4,5,7,8,9: score=2 If respondent answered 10: score=missing
Health Status (for creation of SES)	In general, how would you rate your health today? *1 = Very good; 2 = Good; 3 = Neither poor nor good; 4 = Poor; 5 = Very poor*	Inverted
Received needed medical care	In the last 12 months, has there been a time when you needed health care but did not get that care? *1 = Yes; 2 = No; 3 = No need for health care in the past 12 months*	If respondent answered 1: score=0 If respondent answered 2 or 3: score=1
Involved in making treatment decisions	On your last visit to a health care provider, to what extent were you involved in making decisions for your treatment? *1 (Not at all); 2; 3; 4; 5 (Completely)*	No transformation
Received needed rehabilitation services	In the last 12 months, has there been a time when you needed rehabilitation services, such as physical, occupational, or speech therapy, but did not get those services? *1 = Yes; 2 = No; 3 = No need for rehabilitation services in the past 12 months*	If respondent answered 1: score=0 If respondent answered 2 or 3: score=1
Aware of financial services	Do you know how to get financial services such as credit, insurance, grants, and savings programs? *1 = Yes; 2 = No*	If respondent answered 1: score=1 If respondent answered 2: score=0
Receive social protection	Do you currently benefit from any social protection program, such as loss of income through old age, sickness or disability? *1 = Yes; 2 = No*	If respondent answered 1: score=1 If respondent answered 2: score=0
Participation in self-help group	Are you a member of a self-help group? *1=Yes; 2=No, but I would like to; 3=No, I don't want to*	If respondent answered 1 or 3: score=1 If respondent answered 2: score=0

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract Pg 1, 1-3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found Pg 2, 3-28
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported Pg 4, 2-39
Objectives	3	State specific objectives, including any prespecified hypotheses Pg 4, 40- pg 5, 3
Methods		
Study design	4	Present key elements of study design early in the paper Pg 5, 7-8; 14
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection Pg 5, 14-37
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants Pg 5, 27; pg 8, 5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable Pg 5, 40- pg 6, 23; Supplementary Table
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group Pg 5, 7-13
Bias	9	Describe any efforts to address potential sources of bias Pg 7, 20-27 → The analysis uses PSM, a method to reduce bias
Study size	10	Explain how the study size was arrived at Pg 8, 2-5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why Pg 5, 40- pg 6, 23; Supplementary Table
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding

	Pg 6, 6-29
(*b*) Describe any methods used to examine subgroups and interactions	
(*c*) Explain how missing data were addressed Pg 7, 1-2	
(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
(*e*) Describe any sensitivity analyses Pg 7, 23-27; pg 9, 8-14	

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed Pg 8, 2-4
		(b) Give reasons for non-participation at each stage Pg 8, 7-8
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders Pg 8, 9-10, 21-25
		(b) Indicate number of participants with missing data for each variable of interest Pg 7, 1-2
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures Pg 5, 40- pg 6, 2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included Pg 9, 1-5
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses Pg 9, 5-14
Discussion		
Key results	18	Summarise key results with reference to study objectives Pg 9, 20-26
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias Pg 10, 7-22; 29-33; pg 10, 38- pg 11, 1-7
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence Pg 11, 8-18
Generalisability	21	Discuss the generalisability (external validity) of the study results Pg 11, 4-6
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based Pg 11, 32-33

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.