BMJ Open

Can propensity score matching be applied to cross-sectional data to evaluate Community-Based Rehabilitation? Results of a survey implementing the WHO's Community-Based Rehabilitation indicators in Vietnam

Catherine Mason,[1] Carla Sabariego,[2] Đoàn Mạnh Thắng,[3] Jörg Weber[4]

To cite: Mason C, Sabariego C, Thắng ĐM, *et al*. Can propensity score matching be applied to cross-sectional data to evaluate Community-Based Rehabilitation? Results of a survey implementing the WHO's Community-Based Rehabilitation indicators in Vietnam. *BMJ Open* 2019;**9**:e022544. doi:10.1136/bmjopen-2018-022544

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-022544).

Received 22 February 2018
Revised 7 November 2018
Accepted 14 December 2018

For numbered affiliations see end of article.

Correspondence to
Catherine Mason;
catherine.mason@med.uni-muenchen.de

ABSTRACT

Objectives Community-Based Rehabilitation (CBR) is a multi-sectoral approach working to equalise 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 randomisation and longitudinal data rarely possible. Statistical methods, such as propensity score matching (PSM), offer an alternative approach to evaluate a treatment when randomisation 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 WHO'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 10 covariates. In the unmatched sample, significant differences between groups were found for six of the 10 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 to 19.32) than non-CBR participants (mean=20.93, SD=6.16, 95% CI 19.50 to 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 analysing 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-randomised cross-sectional data, implying the need for statistical approaches, such as PSM, to account for these limitations.
- ► PSM attempts to mimic randomisation 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 analysing cross-sectional CBR data was demonstrated, as biases detected in the distribution of covariates between groups in the unmatched sample 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 equalise 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 synthesise global perspectives on CBR, the WHO developed 'Community-Based Rehabilitation Guidelines' in 2010, which have since become accepted as a conceptual framework for CBR.[2] With these guidelines, WHO emphasised 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 standardised approach to do this.[3 4]

The complexity of CBR leads to challenges in research when attempting to quantify its effectiveness.[5–7] Fully experimental studies with randomisation 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 randomisation 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 randomisation 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 data set 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 randomised controlled studies (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 randomisation 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 alignment 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 programme. 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 organisations' (NGO) activities; and Hòa Bình, where CBR is not implemented by either government or NGOs. The Huế CBR programme began in 2009 in cooperation with the Huế Rehabilitation Hospital. The programme 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 programme 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 centres. 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: https://play.google.com/store/apps/details?id=com.universaltools.whocbr-survey&hl=en).

People with disabilities were identified prior to the survey by government records. In both provinces a team of five local healthcare workers were trained by the lead researcher (CM) over 2 days on how to conduct interviews

Table 1 WHO CBR Indicators and questions used to measure them

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 who 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?
Empowerment	**% 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?

Base indicators are shown in bold. The response option for all questions ranged from 1 (Not at all) to 5 (Completely).

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 the respondent from being interviewed, or if the respondent was a minor, a proxy interview with a family member was performed.

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).

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 dichotomised 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 online supplementary table). Seeing as the CBR programme 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.

Table 2 Baseline characteristics of CBR participants and non-CBR participants in the unmatched and matched samples

Variable	Unmatched Sample No CBR (n=151)	Unmatched Sample With CBR (n=147)	Unmatched Sample Std. dif. of means	Matched Sample No CBR (n=74)	Matched Sample With CBR (n=74)	Matched Sample 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

Absolute standardised differences of means are shown, with differences exceeding the threshold of 0.25 indicated in bold.
Note: continuous variables are presented as means ± SD; dichotomous variables are presented as n(%).

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 callipers of width equal to 0.25 of the 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 standardised 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 standardised 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 two 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 V,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 programme, 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 standardised 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 standardised differences ranged from 0 to 0.147 with all values falling below the threshold value of 0.25[27] (table 2).

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 to 19.32) than non-CBR participants (mean=20.93, SD=6.16, 95% CI 19.50 to 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 standardised 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 to 19.47) than non-CBR participants (mean=21.17, SD=6.381, 95% CI 19.67 to 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 analysing cross-sectional data in the field of CBR. The study analysed 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 programme 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 programme. 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] Similar 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 4 years and 7 years. In our study, data were collected 7 years after the programme 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 emphasise 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 programme being examined. Though CBR aims to impact all aspects of the lives of people with disabilities to increase community inclusion, the programme in Huế does not directly target community inclusion. The programme 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 programme, 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 programme in its early stages, it may be more important to match the indicators used with the specific targets of programmes.

To our knowledge, this study is the first to implement the recently developed WHO CBR Indicators.[4] The study highlights how important it is to collect standardised data in the field of CBR in order to facilitate comparisons between groups and determine effectiveness of programmes. 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,[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 analysing cross-sectional CBR data. Further studies should look into additional statistical methods for analysing 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 standardised 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 emphasise 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 analysing cross-sectional CBR data. While randomised and longitudinal data are ideal for evaluations, cross-sectional data presents the advantage of being more feasible to collect and thereby provides an essential foundation to generate hypotheses and perform further studies.

Therefore, it is essential that appropriate statistical methods are applied to capitalise on available data. The potential of using PSM for analysing 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.

Author affiliations
[1]Department for Medical Information Processing, Biometry and Epidemiology (IBE), Ludwig-Maximilians-Universität München, Munich, Germany
[2]Ludwig-Maximilians University, Public Health and Health Services Research, Munich, Germany
[3]Hoa Binh Department of Health, Hoa Binh, Viet Nam
[4]CBM eV, Bensheim, Hessen, Germany

Acknowledgements 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.

Contributors Conceptualisation: 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.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethical approval was obtained through the Ludwig-Maximilians-Universität Munich Ethics Commission and by the local provincial Ministries of Health.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data are owned by the 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 contacting WHO under disability@who.int.

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:e0178418.
5. Finkenflügel H, Wolffers I, Huijsman R. The evidence base for community-based rehabilitation: a literature review. *Int J Rehabil Res* 2005;28:187–201.
6. Cornielje H, Velema JP, Finkenflügel H. Community based rehabilitation programmes: monitoring and evaluation in order to measure results. *Lepr Rev* 2008;79:36–49.
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.
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;66:1182–6.
9. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
10. Heckman JJ, Ichimura H, Todd P. Matching as an econometric evaluation estimator. *Rev Econ Stud* 1998;65:261–94.
11. Rosenbaum PR. The role of known effects in observational studies. *Biometrics* 1989;45:557–69.
12. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
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;33:1893–901.
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;67:932–9.
15. Li M. Using the propensity score method to estimate causal effects a review and practical guide. *Organ Res Methods* 2013;16:188–226.
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;27:5–18.
17. Grandisson M, Hébert M, Thibeault R. A systematic review on how to conduct evaluations in community-based rehabilitation. *Disabil Rehabil* 2014;36:265–75.
18. Becerril J, Abdulai A. The impact of improved maize varieties on poverty in mexico: A propensity score-matching approach. *World Dev* 2010;38:1024–35.
19. Jalan J, Ravallion M. Estimating the benefit incidence of an antipoverty program by propensity-score matching. *Journal of Business & Economic Statistics* 2003;21:19–30.
20. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci* 2010;25:1–21.
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;13:320.
22. Pollack CE, Chideya S, Cubbin C, *et al*. Should health studies measure wealth? A systematic review. *Am J Prev Med* 2007;33:250–64.
23. Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc* 2007;99:1013.
24. Garrido MM, Kelley AS, Paris J, *et al*. Methods for constructing and assessing propensity scores. *Health Serv Res* 2014;49:1701–20.
25. Buuren Svan, Groothuis-Oudshoorn K. mice : Multivariate Imputation by Chained Equations in *R*. *J Stat Softw* 2011;45:1–67.
26. Cochran WG, Rubin DB. Controlling bias in observational studies: A review. *Sankhya Ser A* 1973:417–46.
27. De H, Imai K, King G, *et al*. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal* 2007;15:199–236.
28. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med* 2008;27.12:2037–49.
29. Keele L. An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. R package version 2.1. 2010 https://CRAN.R-project.org/package=rbounds.
30. De H, Imai K, King G, *et al*. Matchit: Nonparametric preprocessing for parametric causal inference. *J Stat Softw* 2011;42–1–28.
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;68:1102–8.
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;36:1508–17.
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;3:951–79.
34. Implementing the vision: Addressing challenges to results-focused management and budgeting. *Implementation Challenges in Results Focused Management and Budgeting*. Paris: OECD, 2002.
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;142:1380–2.
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;39:246–57.