



SPECIAL ARTICLE

COCHRANE REHABILITATION CORNER 5th COCHRANE REHABILITATION METHODOLOGICAL MEETING

Introduction to target trial emulation in rehabilitation: a systematic approach to emulate a randomized controlled trial using observational data

Pierre CÔTÉ¹, Stefano NEGRINI^{2,3}, Sabrina DONZELLI^{4,5}, Carlotte KIEKENS³,
Chiara ARIENTI⁶, Maria G. CERAVOLO⁷, Douglas P. GROSS⁸, Irene BATTEL^{2*},
Giorgio FERRIERO^{9,10}, Stefano G. LAZZARINI⁶, Bernard DAN^{11,12}, Heather M. SHEARER^{1,13},
Jessica J. WONG¹, Participants in the 5th Cochrane Rehabilitation Methodological Meeting ‡

‡ Members are listed at the end of the paper.

¹Institute for Disability and Rehabilitation Research, Faculty of Health Sciences, Ontario Tech University, Oshawa, ON, Canada; ²Department of Biomedical, Surgical and Dentals Sciences, University “La Statale”, Milan, Italy; ³IRCCS Istituto Ortopedico Galeazzi, Milan, Italy; ⁴Department of Orthopedics, University Medical Center, Utrecht, the Netherlands; ⁵Harvard T.H. Chan School of Public Health, Boston, MA, USA; ⁶IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; ⁷Department of Experimental and Clinical Medicine, Polytechnic University of Marche University, Ancona, Italy; ⁸Department of Physical Therapy, University of Alberta, Edmonton, AB, Canada; ⁹Department of Biotechnology and Life Sciences, University of Insubria, Varese, Italy; ¹⁰Physical Rehabilitation Medicine Unit, Scientific Institute of Tradate IRCCS, Istituti Clinici Scientifici Maugeri, Tradate, Varese, Italy; ¹¹Faculty of Psychology and Educational Sciences, Université Libre de Bruxelles, Brussels, Belgium; ¹²Inkendaal Rehabilitation Hospital, Vlezenbeek, Belgium; ¹³Division of Research, Canadian Memorial Chiropractic College, Toronto, ON, Canada

*Corresponding author: Irene Battel, Department of Biomedical, Surgical and Dentals Sciences, University “La Statale”, Via Festa del Perdono 7, 20122 Milan, Italy. E-mail: irene.battel@gmail.com

This is an open access article distributed under the terms of the Creative Commons CC BY-NC-ND license which allows users to copy and distribute the manuscript, as long as this is not done for commercial purposes and further does not permit distribution of the manuscript if it is changed or edited in any way, and as long as the user gives appropriate credits to the original author(s) and the source (with a link to the formal publication through the relevant DOI) and provides a link to the license. Full details on the CC BY-NC-ND 4.0 are available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

ABSTRACT

Rehabilitation providers and policymakers need valid evidence to make informed decisions about the healthcare needs of the population. Whenever possible, these decisions should be informed by randomized controlled trials (RCTs). However, there are circumstances when evidence needs to be generated rapidly, or when RCTs are not ethical or feasible. These situations apply to studying the effects of complex interventions, including rehabilitation as defined by Cochrane Rehabilitation. Therefore, we explore using the target trial emulation framework by Hernán and colleagues to obtain valid estimates of the causal effects of rehabilitation when RCTs cannot be conducted. Target trial emulation is a framework guiding the design and analysis of non-randomized comparative effectiveness studies using observational data, by emulating a hypothetical RCT. In the context of rehabilitation, we outline steps for applying the target trial emulation framework using real world data, highlighting methodological considerations, limitations, potential mitigating strategies, and causal inference and counterfactual theory as foundational principles to estimating causal effects. Overall, we aim to strengthen methodological approaches used to estimate causal effects of rehabilitation when RCTs cannot be conducted.

(Cite this article as: Côté P, Negrini S, Donzelli S, Kiekens C, Arienti C, Ceravolo MG, et al.; Participants in the 5th Cochrane Rehabilitation Methodological Meeting. Introduction to target trial emulation in rehabilitation: a systematic approach to emulate a randomized controlled trial using observational data. Eur J Phys Rehabil Med 2024;60:145-53. DOI: 10.23736/S1973-9087.24.08435-1)

KEY WORDS: Rehabilitation; Comparative effectiveness research; Randomized controlled trial; Cohort studies; Observational study.

Rehabilitation providers need quality evidence to make timely clinical decisions. Similarly, policymakers require valid research to plan and implement healthcare delivery models that address emerging rehabilitation needs in the population. In an ideal world, these decisions should be informed by randomized controlled trials (RCTs). However, in circumstances where RCTs are not warranted or possible (e.g., when effectiveness data is urgently needed, RCTs are not ethical or feasible for various reasons including costs), best practices or observational studies are often used to inform clinical decisions. For example, the delivery of rehabilitation to patients with moderate to severe COVID-19 during the early phases of the pandemic¹⁻³ or the field of stroke rehabilitation⁴ are largely informed by best practices and observational studies. Therefore, it is necessary to consider how alternative designs to RCTs can provide valid estimates of the cause-and-effect relationship between an intervention for rehabilitation and functioning.

Target trial emulation is a framework used to guide the design and analysis of comparative effectiveness (non-randomized) studies using “real world” observational data that emulates a hypothetical RCT. In this article, we explore how applying the target trial emulation framework by Miguel Hernán and colleagues can provide valid estimates of the causal effects of rehabilitation when RCTs cannot be conducted.^{5,6} Furthermore, we outline the methodology used to design cohort studies and quasi-experimental studies based on target trial emulation, including trial protocol development and the use of observational data necessary to make causal inferences. As illustrated below, our discussion focuses on investigating the comparative effectiveness of rehabilitation grounded in causal inference and extends the principles of benchmarking-controlled trials proposed by Malmivaara.⁷

Ethics, scientific justification, feasibility, and timeliness of RCTs: a balancing act

It is indisputable that RCTs are the design of choice to answer causal questions about the effectiveness of an intervention, such as rehabilitation. The methodological advantage of RCTs comes from the ability to control for confounding through randomization. However, randomization is indicated under specific circumstances: 1) when clinical equipoise is present;⁸ 2) when the interventions being tested do not exceed accepted minimal risks; 3) when it is acceptable to patients; or 4) previous research justifies its conduct.⁹

Establishing clinical equipoise is necessary yet challenging and should be demonstrated by conducting a systematic review of the literature, conducting qualitative research or surveying experts, health care providers and patients, and by understanding patient preferences.¹⁰ The ethical conduct of an RCT also requires that the interventions being tested do not exceed accepted minimal risks. For example, it would be unethical to estimate the effects of high-dose opioids on pain intensity in children with cerebral palsy because of the high risk of addiction and other side effects.

There are also situations when conducting an RCT is not feasible because evidence is urgently needed to make clinical timely decisions (e.g., rehabilitation for patients in emergency situations), the costs of conducting an RCT can be prohibitive (e.g., robotic rehabilitation combined with brain-computer interface technology for patients with spinal cord injury) and recruiting enough participants may not be possible because a disease is rare. In rehabilitation, RCTs are often not feasible as interventions in rehabilitation have small effects and the disease may be rare, posing challenges with sample size and reliability. Nonetheless, evidence is strongly needed and strategies with alternative study designs can overcome these challenges. Finally, as described by Shearer *et al.*,⁹ testing the effectiveness of rehabilitation in an RCT requires that preliminary research supports its feasibility, potential benefits, and safety for patients.

Rehabilitation and comparative effectiveness

According to Cochrane Rehabilitation, rehabilitation includes a multimodal, person-centered, and collaborative process, that targets a person’s capacity or contextual factors related to performance with the goal of optimizing functioning.^{9,11} Rehabilitation is inherently complex and needs to be tailored to an individual’s needs. The complexity of rehabilitation resides in its structure, delivery mode and targeted outcomes, while its design requires combining clinical interventions and technologies, clinicians, settings, and recipients of the intervention (patients, family, and community). Therefore, applying the target trial emulation framework to rehabilitation necessitates careful consideration of the complexity of the experimental and control interventions. Since rehabilitation needs to be tailored to the specific needs of study participants, it is necessary that the sources of data used to emulate a trial describe the ingredients of the rehabilitation program in sufficient detail for clinicians to understand what and how it was tailored to patients’ individual needs. This descrip-

TABLE I.—Treatment strategies described by Heil et al.,¹² used in their study to determine the effectiveness of a multimodal prehabilitation program compared with usual care in high-risk patients with colorectal cancer who underwent elective colorectal surgery.

Description of multimodal prehabilitation program	Description of usual care
<ul style="list-style-type: none"> • Multimodal prehabilitation consisted of case management of a specialized oncology nurse and anemia treatment. • In addition, patients were strongly advised to reduce intoxications (smoking cessation and reduction of alcohol intake). • During intake with the physical therapist, a personalized exercise program was made for each participant. • Each patient received tailored nutritional advice from a dietician. The exercise program designed by the physical therapist contained two components: 1) three times a week, for at least 3 weeks, a 60 min high-intensity training in the hospital supervised by a physical therapist; 2) four times a week for at least 60 min, a non-supervised low-intensity endurance training at home (e.g., walking or biking). • During the intake with a dietician, patients received tailored nutritional advice to achieve a total protein intake of 1.9 g per kg of lean body mass per day. Patients were also advised to take an additional 0.4 g per kg protein, within 1 h before high-intensity training and daily before bedtime. If necessary, protein shakes were prescribed to achieve this intake. • During prehabilitation, patients were followed by a dietician with a final consultation at the end of the program. 	<ul style="list-style-type: none"> • Anemia treatment as indicated (using intravenous iron medication or blood transfusion per protocol) • A 30-minute preoperative assessment with the physical therapist for breathing exercises • A preoperative calculation of the nutritional assessment score (SNAQ). In case the SNAQ-score ≥ 3, patients were referred for a consultation with a dietician. • Patients were accompanied from diagnosis to the end of the treatment trajectory (including both surgery and chemotherapy) by a specialized oncology nurse who provided detailed information about the diagnosis and treatment process, as well as psychological support. • On indication, a psychologist was consulted

tion requires that the data about the intervention are sufficiently detailed to differentiate the experimental from the control group and allow the valid classification of patients into the correct intervention group.

It is useful to consider the emulated target trial by Heil *et al.* to illustrate how rehabilitation compared within a target trial intervention framework should be described.¹² The authors aimed to determine the effectiveness of a multimodal prehabilitation program compared with usual care in high-risk patients with colorectal cancer who underwent elective colorectal surgery. The detailed treatment strategies (multimodal prehabilitation *versus* usual care) as described by Heil *et al.*¹² are outlined in Table I. Heil *et al.* provided adequate details of the intervention and comparison groups, including all components and provider details of the multimodal prehabilitation program and usual care, respectively.¹² These details were captured in the data source they used and allowed for correct classification of patients into the two groups when emulating a target trial.

Causal inference

Making valid causal inferences requires that we use counterfactual theory.¹³⁻¹⁵ In an ideal situation, we would compare the potential outcomes for the same individual under two contrasting conditions (e.g., treatment A *versus* treatment B) for causal inference. However, one of these conditions is real (factual) and the outcome can be observed, while the other condition is counterfactual and the outcome cannot be observed (termed ‘counterfactual’ or

‘potential outcome’). Let us apply counterfactual theory to studying whether a vocational counselling intervention is associated with return to work in patients with rheumatoid arthritis. To determine whether vocational counselling improved the return-to-work rate requires that we answer the question “Would the patients have returned to work at the same time (outcome) had they not received vocational counselling?”. In an ideal situation, we would be able to observe the outcome for the same patients at the same time but treated with usual care instead of vocational counselling. Of course, this is not possible, and we cannot make this observation; it is counterfactual. Since observing the outcome in the same individuals at the same time under two distinct intervention scenarios is impossible, we must rely on comparative experiments to estimate the effect of an intervention.

To make causal inferences using counterfactual theory, a study must meet the exchangeability, or “no confounding and no selection bias” assumption. Exchangeability implies that the groups being compared (exposed and unexposed) are similar in all aspects that are of importance to the outcome, except for the exposure.^{16, 17} Achieving exchangeability is straightforward in randomized controlled trials but more challenging in observational studies. In observational studies, achieving exchangeability hinges on adequately controlling for confounding, including unmeasured factors through design (e.g., self-controlled studies, Directed Acyclic Graphs (DAGs) and negative control exposures) or analysis (e.g., instrumental variables or model-based adjustment for confounders).^{18, 19}

Applying the target trial emulation framework to quantify the effect of rehabilitation

As described earlier, it is impossible to observe the outcome in the same individual at the same time under two contrasting treatment conditions. To overcome this fundamental problem, we can move to estimating population causal effects, which involves comparing outcomes under different conditions between comparable groups. Tied to the previous example, to quantify the effect of vocational counselling (as an example of rehabilitation) on return to work in patients with rheumatoid arthritis, the rate of return to work can be compared between groups of patients who received vocational counselling to those who did not receive vocational counselling (but received usual care instead). Ideally, these patients with rheumatoid arthritis would be randomly assigned to the two groups to ensure that they are comparable (exchangeable) at baseline. Therefore, any differences in return-to-work rates can be attributed to vocational counselling as opposed to prognostic differences between the two groups. RCTs also provide the advantage of clearly specifying time zero for each participant (time of randomization to the assigned treatment group), which is important for causal inference.

If an RCT cannot be conducted due to feasibility, ethical issues, or time/resource constraints, observational studies can be designed to incorporate these key features of an RCT for causal inference. Applying the target trial emulation framework, developed by Hernan and colleagues, facilitates this by explicitly emulating a hypothetical RCT that is designed to answer the causal question at hand, which is known as the ‘target trial’.^{5, 6, 20, 21} Suppose we would like to design a study to determine the effect of dedicated units of inpatient stroke rehabilitation (a specific model of service delivery in Ontario, Canada) on the risk of fractures in adults with a diagnosis of stroke or transient ischemic attack (TIA). Regarding this specific model of service delivery, dedicated units for inpatient stroke rehabilitation are units in rehabilitation facilities with beds and therapists dedicated to patients with stroke (Table II).²² There are potential chal-

lenges to conducting an RCT comparing dedicated *versus* non-dedicated units for inpatient stroke rehabilitation, as these units represent complex interventions rolled out to different facilities at different times, and an RCT with a 2-year follow-up to assess the risk of stroke would be time- and resource-intensive. If we are unable to conduct an RCT to answer this causal question, we can consider emulating a target trial with analyses of observational data. Target trial emulation involves two steps:

Step 1: Articulating the causal question and protocol

The causal question and protocol would be formulated as though we were designing a hypothetical RCT. This causal question would be the question we ask if we could design an RCT *e.g.*, What is the effect of inpatient stroke rehabilitation delivered through dedicated facility units compared to inpatient stroke rehabilitation delivered through non-dedicated stroke rehabilitation units on risk of fracture (within 2 years post-stroke) in older adults with a diagnosis of stroke? (Table II). The protocol outlines key elements (termed “causal estimands”) of a randomized trial representing the target study to answer this causal question. This includes eligibility criteria, treatment strategies, treatment assignment, start and end of follow-up, outcomes, causal contrasts, and analysis plan^{5, 6} (Table III).

Target trial protocol for our example

Using the example of inpatient stroke rehabilitation, we would start by designing a comparative effectiveness study with clear eligibility criteria for the sample (component #1), *e.g.*, older adults (aged ≥ 65 years) with a diagnosis of stroke or TIA, with specific inclusion and exclusion criteria listed in Table III. Component #2 describes the treatment strategies. For example, this can be dedicated units *versus* non-dedicated units of inpatient stroke rehabilitation.²² Component #3 describes how participants are assigned to a treatment strategy in this hypothetical RCT, *e.g.*, randomly assigned to dedicated *versus* non-dedicated stroke rehabilitation units at baseline in this target RCT.

TABLE II.—Example of causal question related to rehabilitation using Population, Intervention, Comparison, Outcome and Time (PICOT) framework.

PICOT Component	Example
Population	Older adults (aged ≥ 65 years) with a diagnosis of stroke
Intervention	Dedicated units for inpatient stroke rehabilitation (units with geographically distinct, stroke-dedicated beds and dedicated therapists) ²²
Comparison	Non-dedicated units for inpatient stroke rehabilitation
Outcome	Risk of fracture
Time (of follow-up)	2 years after baseline (<i>i.e.</i> , after index date of stroke)

TABLE III.—Components of protocol using the target trial emulation framework^{5, 6, 20, 21} applied to estimating the effect of a type of rehabilitation unit on the risk of fracture in older adults with stroke.

#	Protocol component	Specification of target trial	Emulation of target trial using observational data
1	Eligibility criteria	<ul style="list-style-type: none"> • Inclusion criteria: Older adults (aged ≥ 65 years) with a diagnosis of stroke or transient ischemic attack • Exclusion criteria: Adults with subarachnoid hemorrhage or with preexisting hemiparesis (for those with transient ischemic attack) 	<ul style="list-style-type: none"> • Same as specification • Required in observational data: age, diagnosis of stroke or transient ischemic attack, history of subarachnoid hemorrhage, pre-existing hemiparesis
2	Treatment strategies	<ul style="list-style-type: none"> • Dedicated units for inpatient stroke rehabilitation (units with geographically distinct, stroke-dedicated beds and dedicated therapists).²² This represents one model of service delivery in Ontario • Non-dedicated units for inpatient stroke rehabilitation, representing a different model of service delivery 	<ul style="list-style-type: none"> • Same as specification • Required in observational data: dedicated <i>versus</i> non-dedicated units for inpatient stroke rehabilitation
3	Assignment procedures	<ul style="list-style-type: none"> • Participants will be randomly assigned to dedicated units of inpatient stroke rehabilitation <i>or</i> non-dedicated units of inpatient stroke rehabilitation at baseline • Participants will be aware of treatment assignment (cannot be blinded due to the nature of the intervention) 	<ul style="list-style-type: none"> • Different from specification (<i>i.e.</i>, no randomization) • Eligible persons will be assigned to dedicated <i>versus</i> non-dedicated units of inpatient stroke rehabilitation based on what was documented in observational data at the time of eligibility
4	Follow-up period	<ul style="list-style-type: none"> • Begins at randomization (baseline) • Ends at fracture, loss to follow-up, death, or 2 years after baseline 	<ul style="list-style-type: none"> • Same as specification • Required in observational data: date of fracture, date of loss to follow-up, date of death
5	Outcome	<ul style="list-style-type: none"> • Low-trauma fracture that occurred within 2 years of index stroke, defined as any fracture of the femur, forearm, humerus, pelvis, or vertebrae and excludes fractures from trauma, traffic collisions, falls from a height, or in people with active cancer⁶ 	<ul style="list-style-type: none"> • Same as specification • Required in the observational data source: Fracture, including type of fracture, cancer data
6	Causal contrasts of interest	<ul style="list-style-type: none"> • Intention-to-treat effect • Per-protocol effect 	<ul style="list-style-type: none"> • Observational analogue of intention-to-treat effect
7	Analysis plan	<ul style="list-style-type: none"> • Comparison of fracture risk among participants assigned to dedicated units <i>versus</i> non-dedicated units of inpatient stroke rehabilitation • Intention-to-treat or per-protocol analysis • All analyses adjust for baseline confounders/prognostic factors 	<ul style="list-style-type: none"> • Analyses adjusting for baseline confounders/prognostic factors • Required in observational data: All baseline confounders/prognostic factors

Component #4 describes the follow-up period from baseline (time zero, starts at randomization) to the outcome or end of follow-up, such as ending at the outcome (fracture), death, loss to follow-up, or 2 years after baseline. Component #5 states the outcome of interest, *e.g.*, risk of fracture within 2 years after baseline (*i.e.*, after index date of stroke or TIA). Component #6 refers to causal contrasts such as intention-to-treat effect or per-protocol effect. Lastly, component #7 describes the analysis plan, which would emulate the analysis intended for the target trial, such as to estimate the intention-to-treat or per-protocol effect.

Step 2: Execution of target trial emulation - Emulating the components of a protocol representing the target trial

The protocol components representing a target trial (hypothetical randomized trial) would be emulated using observational data (Table III). This hypothetical randomized

trial represents the target study for causal inference. This step involves adapting the target trial protocol to using observational data, which includes finding eligible individuals, assigning eligible individuals to a treatment strategy based on what is documented in the observational data, following individuals from time zero (time of treatment assignment) up to the outcome or end of follow-up, and developing an analytic plan aligned with the target trial with the addition of adjusting for baseline confounders to emulate randomization. To guide the study design and analysis, directed acyclic graphs (DAGs) are used to represent relationships between variables, including the exposure or treatment assignment, the outcome, and confounders.^{18, 19} This guides the conceptual thinking on whether certain variables are confounders, including measured and unmeasured confounders. Emulating the target trial requires that the necessary data elements are captured in the observational data source, as outlined in Table III. For

this example, we needed data on the diagnosis of stroke (population), type of inpatient stroke rehabilitation (treatment strategy), fracture risk (outcome), and a wide range of prognostic factors (confounders), in addition to other data elements for the eligibility criteria and lost to follow-up. This represents the necessary data requirement to execute target trial emulation. Therefore, when designing a comparative effectiveness study with observational data for target trial emulation, we need to map the observational data and analysis onto the specifications of the target trial. However, due to data constraints, we may need to modify the protocol components, such as eligibility criteria, treatment strategies, outcome, or follow-up period. We aim to design a target trial as close to the ideal trial as possible based on available observational data by outlining the protocol, describing how observational data is used to emulate the target trial, and explaining how the target trial differs from the ideal trial.^{6, 23-25}

Regarding the analytic plan using observational data, we aim to examine the intention-to-treat effect, which compares the outcomes of groups assigned to each treatment at baseline, where possible based on the observational data available.^{5, 6, 20, 21} This includes comparing the initiation of different treatments (e.g., data on the prescription of different drugs) as an observational analog of the intention-to-treat analysis. Baseline confounders are accounted for in the analysis in the absence of randomization to a treatment strategy. In our example, the analytic plan includes adjustment of baseline confounders/prognostic factors, and potential confounders of the association between service delivery of inpatient rehabilitation and risk of fracture include age, sex, previous fractures and health-related conditions, medication use, and stroke-specific fac-

tors including hemiplegia/hemiparesis, visual impairment, stroke type and severity, cognitive function, and functional status after stroke.^{22, 26}

Use of observational data for target trial emulation

Target trial emulation typically involves analysis of observational data from large databases, such as disease registries, population health surveys, administrative data, or potentially multiple data sources that are linked. This is because the observational data needs to capture information aligned with the eligibility criteria, treatment strategy, outcome, and a wide range of potential confounders to emulate the target trial. This is the necessary data requirement to execute target trial emulation for our example of inpatient stroke rehabilitation (Table IV provides an example of these data sources). For example, if observational data sources do not measure the components of the rehabilitation program or the outcome of interest, we would not be able to emulate the target trial. Similarly, if the observational data sources do not capture information on a wide range of potential confounders, there is the risk of residual confounding from unmeasured confounders and the emulated study would not provide valid estimates of effects. If applicable, it is important to consider the feasibility of linking multiple databases if needed to conduct target trial emulation for the research at hand, as there may be ethical or administrative barriers to data linkages. Our target trial example focused on inpatient stroke rehabilitation proposes the use of multiple large registries and administrative databases that can be linked at ICES (originally known as the Institute for Clinical Evaluative Sciences) (Table IV).

TABLE IV.—Description of linked data holdings at ICES for target trial emulation in the stroke rehabilitation example.

Description of ICES (originally known as the Institute for Clinical Evaluative Sciences)	Examples of data holdings at ICES used for stroke rehabilitation example
<ul style="list-style-type: none"> • In the stroke rehabilitation example of this paper, answering the proposed causal question using target trial emulation could be conducted using data holdings at ICES, including linked registry databases and population-based administrative data • ICES is an independent, non-profit research institute whose legal status under Ontario’s health information privacy law allows it to collect and analyze health care and demographic data, without consent, for health system evaluation and improvement. • Data holdings at ICES, including linked registry databases and population-based administrative data, could be used for Target Trial Emulation of the stated example because they capture relevant information on inpatient stroke rehabilitation, risk of fracture, a wide range of potential confounders, and other data elements as described in Table III 	<ul style="list-style-type: none"> • Ontario Stroke Registry, which collects clinical information on a population-based sample of patients with stroke or TIA seen at all acute care institutions in Ontario • National Rehabilitation Reporting System (NRS), which collects data on inpatient rehabilitation, including rehabilitation services for stroke and TIA <p>Ontario Stroke Registry and NRS can be linked to administrative databases using unique, encoded patient identifiers:</p> <ul style="list-style-type: none"> • Canadian Institute for Health Information Discharge Abstract Database (CIHI-DAD) and National Ambulatory Care Reporting System (NACRS) to capture hospitalizations and emergency department visits for fractures, falls, and other medical conditions • Ontario Health Insurance Plan (physician claims) database to capture outpatient physician visits and procedures • Registered Persons Database (RPDB) to identify deaths • Ontario Drug Benefits database to capture medication claims (for those ≥65 years)

Overall, the target trial emulation framework guides the design of comparative effectiveness studies using observational data to emulate a protocol that would be used for the target trial (randomized trial). In the example above, potential challenges to conducting this RCT on examining the effects of dedicated units of inpatient stroke rehabilitation on the risk of fracture include the use of rehabilitation process rolled out in different units over time, feasibility, and resources concerning 2-year follow-up for assessing the risk of fracture. To inform timely decisions on types of inpatient stroke rehabilitation, we can consider designing comparative effectiveness studies that emulate the target trial causal question and adapt the target trial protocol using observational data. Therefore, this framework can be used to apply principles of causal inference to estimate the effects of rehabilitation when a randomized trial cannot be conducted, such as because of feasibility, ethical reasons, or time/resource constraints.

Limitations of target trial emulation

To emulate randomization to the intervention *versus* comparison, we need to adjust for confounders to ensure comparability of the two groups at baseline. The selection of confounders should be informed by DAGs.^{18, 19} Various methods can be used to control for confounders, including matching (e.g., propensity score matching), stratification or regression, standardization or inverse probability weighting, g-estimation, or doubly robust methods, as further described by Hernán and Robins.⁶ It is not possible to emulate the target trial, particularly random assignment, if the observational data used does not measure a wide range of potential confounders or if we do not adjust for identified confounders in the analysis plan. Therefore, the observational data source(s) requires measurement of the treatment strategies, outcome (including any adverse effects if possible), and potential confounders.

Example of advanced analytic method to achieve comparability of groups at baseline: propensity score matching

As an example, propensity score matching can be used to account for a wide range of potential confounders in aims to achieve comparability of two groups at baseline. The propensity score is the probability of treatment assignment (e.g., receiving treatment *versus* no treatment) conditional on observed baseline characteristics.²⁷ The propensity score is a balancing score, whereby conditional on the propensity score, the distribution of observed baseline covariates is similar between treatment groups (treated

versus untreated).²⁷ This involves conducting a logistic regression model that includes the potential confounders as independent variables and treatment assignment (e.g., treated *versus* untreated) as the dependent variable to compute and output the propensity score. Subjects between the two treatment groups can be matched on this propensity score, and balance diagnostics assessed to determine whether there is sufficient balance in baseline covariates (<10% standardized mean difference).²⁸ The propensity-score matched cohort can then be used to assess for differences in the outcome(s). Details on propensity score methods can be found in Austin's introduction to propensity score methods,²⁷ which provides other resources on these methods.

It is important to acknowledge that residual confounding from unmeasured confounders remains a limitation of a comparative effectiveness study using target trial emulation. Nevertheless, if there is unmeasured confounding (i.e., confounder not captured in the observational data source), methods can be used to estimate its potential impact, with various methods further described by Lash *et al.*^{29, 30} and VanderWeele *et al.*^{31, 32} As one approach, quantitative bias analyses can be conducted as a sensitivity analysis to assess the potential impact of residual confounding from unmeasured confounders.^{31, 32} Among various methods within quantitative bias analyses, an approach by VanderWeele *et al.* assesses how robust an association is to potential unmeasured confounding through a measure called the E-value.³² E-value is defined as the minimum strength of association (on the risk ratio scale) that an unmeasured confounder would need to have with both the exposure and outcome to fully explain away the observed exposure-outcome association, conditional on measured covariates.³² A large E-value suggests that considerable unmeasured confounding (i.e., an unmeasured confounder strongly associated with both the treatment and outcome) would be needed to explain away an effect estimate.³¹ This allows researchers and others to consider how robust treatment-outcome associations are by assessing whether confounder associations of that magnitude are likely plausible.³²

Calculation of E-value in quantitative bias analysis

For an observed risk ratio (RR), the E-value calculation is:³²

$$\text{E-value} = \text{RR} + \sqrt{[\text{RR} \times (\text{RR} - 1)]}.$$

The mathematical proof for this E-value formula is described elsewhere.³³ This formula applies to an $\text{RR} > 1$, so for $\text{RR} < 1$, the inverse of the observed RR is used in the

TABLE V.—Example of E-value calculation and interpretation in quantitative bias analysis.^{31, 32}

Example of E-value calculation	Description
Observed treatment-outcome association from primary analysis (after adjusting for measured confounders)	Risk ratio (RR) = 3.8
E-value calculation using formula	$E\text{-value} = RR + \sqrt{[RR \times (RR - 1)]} = 3.8 + \sqrt{[3.8 \times (3.8 - 1)]} = 7.1$
Interpretation	<ul style="list-style-type: none"> • An unmeasured confounder would need to be associated with both the treatment and outcome by a risk ratio of 7.1-fold each (above and beyond measured confounders) to fully explain away the observed risk ratio of RR=3.8 • An unmeasured confounder weakly associated with treatment and outcome (<i>i.e.</i>, less than a risk ratio of 7.1-fold each) would not be enough to fully explain away the observed risk ratio of RR=3.8 • Therefore, considerably strong confounding associations (with treatment and outcome) would be needed to fully explain away the observed treatment-outcome association of RR=3.8

formula. See Table V for an example of calculating and interpreting an E-value.

Other limitations include that target trial emulation using observational data cannot be used to study new treatments not yet used in practice and not captured in observational data sources. Routine practice also does not use placebo/sham or blinding of outcome assessment; thus, studies using target trial emulation in this area emulate pragmatic RCTs of rehabilitation. This aligns with rehabilitation that tends to not have a valid placebo/sham due to the nature of the intervention.

Conclusions

This work describes the use of target trial emulation by Hernán and colleagues as applied to rehabilitation. This approach is important to obtain valid estimates of the causal effects of complex interventions, including rehabilitation, because there are circumstances when evidence needs to be generated rapidly, or when RCTs are not ethical or feasible. While there are limitations to this framework as described in this paper, target trial emulation serves as a methodologically robust option to RCTs. Overall, we aim to strengthen methodological approaches used to estimate causal effects of rehabilitation when RCTs cannot be conducted.

References

1. Dillen H, Bekkering G, Gijbbers S, Vande Weygaerde Y, Van Herck M, Haesevoets S, *et al.* Clinical effectiveness of rehabilitation in ambulatory care for patients with persisting symptoms after COVID-19: a systematic review. *BMC Infect Dis* 2023;23:419.
2. Metz J, McElheny K, Robinson JN, Scott DA, Sutton KM, Toresdahl BG. Considerations for Return to Exercise Following Mild-to-Moderate COVID-19 in the Recreational Athlete. *HSS J* 2020;16(Suppl 1):102–7.
3. World Health Organization. Living guidance for clinical management of COVID-19; 2021 [Internet]. Available from: <https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-2> [cited 2024, Jan 29].

4. National Clinical Guideline Centre (UK). *Stroke Rehabilitation: Long Term Rehabilitation After Stroke*. London: Royal College of Physicians (UK); 2013.
5. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* 2022;328:2446–7.
6. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016;183:758–64.
7. Malmivaara A. Benchmarking Controlled Trial—a novel concept covering all observational effectiveness studies. *Ann Med* 2015;47:332–40.
8. Nix HP, Weijer C. Uses of equipoise in discussions of the ethics of randomized controlled trials of COVID-19 therapies. *BMC Med Ethics* 2021;22:143.
9. Negrini S, Kiekens C, Levack WM, Meyer-Feil T, Arienti C, Côté P; Participants in the 5th Cochrane Rehabilitation Methodological Meeting. Improving the quality of evidence production in rehabilitation. Results of the 5th Cochrane Rehabilitation Methodological Meeting. *Eur J Phys Rehabil Med* 2023.
10. Alexander JH. Equipoise in Clinical Trials: Enough Uncertainty in Whose Opinion? *Circulation* 2022;145:943–5.
11. Negrini S, Selb M, Kiekens C, Todhunter-Brown A, Arienti C, Stucki G, *et al.*; 3 Cochrane Rehabilitation Methodology Meeting participants. Rehabilitation Definition for Research Purposes: A Global Stakeholders’ Initiative by Cochrane Rehabilitation. *Am J Phys Med Rehabil* 2022;101:e100–7.
12. Heil TC, Verdaasdonk EG, Maas HA, van Munster BC, Rikkert MG, de Wilt JH, *et al.* Improved Postoperative Outcomes after Prehabilitation for Colorectal Cancer Surgery in Older Patients: An Emulated Target Trial. *Ann Surg Oncol* 2023;30:244–54.
13. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974;66:688.
14. Hernán MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* 2004;58:265–71.
15. Hernan MA, Robins JM. *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC; 2020.
16. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. *Modern Epidemiology*. Wolters Kluwer; 2021.
17. Igelström E, Craig P, Lewsey J, Lynch J, Pearce A, Katikireddi SV. Causal inference and effect estimation using observational data. *J Epidemiol Community Health* 2022;76:960–6.
18. VanderWeele TJ, Hernán MA, Robins JM. Causal directed acyclic graphs and the direction of unmeasured confounding bias. *Epidemiology* 2008;19:720–8.
19. Lipsky AM, Greenland S. Causal Directed Acyclic Graphs. *JAMA* 2022;327:1083–4.
20. Hernán MA. With great data comes great responsibility: publishing comparative effectiveness research in epidemiology. *Epidemiology* 2011;22:290–1.

21. Hernán MA. Methods of Public Health Research - Strengthening Causal Inference from Observational Data. *N Engl J Med* 2021;385:1345–8.
22. Foley N, Meyer M, Salter K, Bayley M, Hall R, Liu Y, *et al*. Inpatient stroke rehabilitation in Ontario: are dedicated units better? *Int J Stroke* 2013;8:430–5.
23. Zhang Y, Thamer M, Kaufman J, Cotter D, Hernán MA. Comparative effectiveness of two anemia management strategies for complex elderly dialysis patients. *Med Care* 2014;52(Suppl 3):S132–9.
24. Garcia-Albeniz X, Chan JM, Paciorek A, Logan RW, Kenfield SA, Cooperberg MR, *et al*. Immediate versus deferred initiation of androgen deprivation therapy in prostate cancer patients with PSA-only relapse. An observational follow-up study. *Eur J Cancer* 2015;51:817–24.
25. Zuo H, Yu L, Campbell SM, Yamamoto SS, Yuan Y. The implementation of target trial emulation for causal inference: a scoping review. *J Clin Epidemiol* 2023;162:29–37.
26. Kapral MK, Fang J, Alibhai SM, Cram P, Cheung AM, Casaubon LK, *et al*. Risk of fractures after stroke: Results from the Ontario Stroke Registry. *Neurology* 2017;88:57–64.
27. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res* 2011;46:399–424.
28. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083–107.
29. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Springer Science & Business Media; 2011.
30. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol* 2014;43:1969–85.
31. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology* 2011;22:42–52.
32. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: introducing the E-Value. *Ann Intern Med* 2017;167:268–74.
33. Ding P, VanderWeele TJ. Sensitivity Analysis Without Assumptions. *Epidemiology* 2016;27:368–77.

Conflicts of interest

The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Funding

This study was supported and funded by the Italian Ministry of Health - Ricerca Corrente 2023. In addition, Pierre Côté's contribution to the project was supported by the Canada Research Chairs Program and Jessica Wong was supported by a Banting Postdoctoral Fellowship from the Canadian Institute of Health Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Pierre Côté and Jessica J. Wong have given substantial contributions to the conception or the design of the manuscript. Stefano Negrini, Sabrina Donzelli, Charlotte Kiekens, Chiara Arienti, Maria G. Ceravolo, Douglas P. Gross, Irene Battel, Giorgio Ferriero, Stefano G. Lazzarini, Bernard Dan, and Heather M. Shearer to the acquisition, analysis and interpretation of the data. All authors have participated in drafting and revisiting the manuscript critically. All authors read and approved the final version of the manuscript.

Group author members

Christopher COLVIN (Charlottesville, VA, USA); Claudio CORDANI (Milan, Italy); Anne CUSICK (Sydney, Australia); Matteo J. DEL FURIA (Milan, Italy); Susanna EVERY-PALMER (Wellington, New Zealand); Peter FEYS (Hasselt, Belgium); Christoph GUTENBRUNNER (Hannover, Germany); Carsten B. JUHL (Copenhagen, Denmark); William M. LEVACK (Wellington, New Zealand); Wendy MACHALICEK (Eugene, OR, USA); Rachelle MARTIN (Wellington, New Zealand); Federico MERLO (Milan, Italy); Thorsten MEYER-FEIL (Halle, Germany); Luca MIRANDA (Milan, Italy); Bianca MOSCONI (Milan, Italy); Randolph NUDO (Kansas City, KS, USA); Aydan ORAL (Istanbul, Türkiye); Cecilie RØE (Oslo, Norway).

History

Manuscript accepted: January 26, 2024. - Manuscript received: January 22, 2024.