A guideline-implementation intervention to improve the management of low back pain in primary care: A difference-in-difference-in-differences analysis

Supplementary Material

Applied Health Economics and Health Policy

Ross Wilson*, Yana Pryymachenko, J. Haxby Abbott, Sarah Dean, James Stanley, Sue Garrett, Fiona Mathieson, Anthony Dowell, Ben Darlow

* Corresponding author:
Centre for Musculoskeletal Outcomes Research
Department of Surgical Sciences
University of Otago
Dunedin, New Zealand
Email: ross.wilson@otago.ac.nz

Funding: The RCT and the implementation study providing data for this study were funded by the Accident Compensation Corporation (https://www.acc.co.nz). BD, JS, SD, JHA, FM, and TD were named investigators on the funding application. For this study, the funder extracted the data from routine insurance claims records, but had no other role in study design, conduct, analysis, interpretation, reporting, or the decision to submit for publication.

SD is partially supported by the National Institute for Health Research Applied Research Collaboration South West Peninsula, UK. The views expressed in this publication are those of the authors and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care, UK.

Conflicts of Interest: The authors declare that they have no conflict of interest

Appendix A: Details of the estimation strategy

Data

All analysis was conducted at the GP-calendar month level (and data for GPs not trained in the FREE approach were extracted at this level). For FREE-trained GPs, data were extracted from ACC's claims database by month relative to the date of training. To form the combined dataset for analysis, these were mapped to the nearest calendar month: for training dates on or before the 15th of the month, the first month's data were matched to the calendar month of training; for training dates on or after the 16th of the month, the first post-training month was matched to the following calendar month (e.g., for a GP trained on 25 August 2017, the first post-training month, representing claims lodged between 26 August and 25 September, was aligned with calendar month data for non-trained GPs for September 2017).

Statistical analysis

We estimated the effects of training using a generalised linear model (GLM) with a log link and quasi-Poisson (for count outcomes) or gamma (for cost data) variance function. (These are exactly equivalent to the Poisson [1,2] and gamma [3] pseudo-maximum likelihood estimators [4].) Our estimated equations take the following form:

$$E\left[y_{ijkt}\right] = \exp\left(\gamma_{kt} + \lambda_{jt} + \theta_{jk} + \beta \mathrm{D}_{ikt}\right),$$

where y_{ijkt} is the outcome for GP i, cohort $j \in \{\text{FREE-trained, non-FREE-trained}\}$, injury type $k \in \{\text{LBP, non-Musuloskeletal}\}$, and time t in months from July 2013 to December 2020. This specification provides a full set of non-parametric controls for injury type-specific time effects γ_{kt} (common across FREE-trained and non-FREE-trained GPs), cohort-specific time effects λ_{jt} (common across LBP and non-musculoskeletal claims), and time-invariant injury type-specific cohort differences θ_{jk} [5]. D_{ikt} is a 'treatment dummy' indicating FREE-trained GPs' LBP claims outcomes in the post-training period; the corresponding coefficient β gives the estimated treatment effect. No other GP characteristics were included in the model (as no such data were available for the non-FREE-trained national GP cohort). The model is estimated separately for each outcome of interest (described in Section 2.2 in the main manuscript). All models were estimated with standard errors clustered at the GP level.

This formulation provides estimates of the average effect of FREE training over the full follow-up period. To further analyse the persistence of any effects of training, we also estimated models where we interacted the treatment dummy D with different functions of the time since training. Specifically, we calculated summary effect estimates for 6-month periods (months 1–6, months 7–12, etc.) post-training and, for graphical presentation of effect trajectories, the full set of month-specific effects (months 1, 2, …, 33 post-training), as well as a natural cubic spline smoothing to visualise the trajectories.

In all analyses, GP-months in which no ACC claims were lodged (for any injury type) were dropped, as these provide no information on the LBP/non-musculoskeletal injury contrast. (Results including all GP-month observations, available from the authors on request, are virtually unchanged from those reported in this article.)

The advantage of the GLM specification is that it is consistent under the weak assumption that the conditional mean function is correctly specified (even if the variance function $\mathrm{Var}(y) \propto \mathrm{E}[y]$ is misspecified [4,6]). The choice of link function captures the assumed multiplicative structure of the data – effects of the training are expected to be proportional to the GP's overall level of activity (captured by the non-musculoskeletal claims data) and prior level of LBP claims activity.

To aid interpretation of the non-linear GLM results, we also report average treatment effects (on the treated; ATT) alongside our coefficient estimates. Average treatment effects were calculated, based on the potential outcomes framework, as the expected difference, per FREE-trained GP over the 33-month follow-up period, between observed (i.e., treated) LBP claims/healthcare visits/costs and predicted outcomes if they had not been treated (where the latter are derived from the assumption of parallel trends (on the log scale)) [7]:

$$\sum_{t \in \{\text{Post-training}\}} E[y(1) \mid j = \text{FREE-trained}, k = \text{LBP}, t] - E[y(0) \mid j = \text{FREE-trained}, k = \text{LBP}, t],$$

where

$$E[y(1) \mid\ j = \text{FREE-trained}, k = \text{LBP}, t] =\ \exp(\gamma_{\text{FREE},t} + \lambda_{\text{LBP},t} + \theta_{\text{FREE,LBP}} + \beta), \text{ and}$$
$$E[y(0) \mid\ j = \text{FREE-trained}, k = \text{LBP}, t] =\ \exp(\gamma_{\text{FREE},t} + \lambda_{\text{LBP},t} + \theta_{\text{FREE,LBP}}).$$

Confidence intervals (at the 95% level) were calculated using the delta method [8].

References

1. Cameron AC, Trivedi PK. Regression analysis of count data. 2nd ed. Cambridge, UK: Cambridge University Press; 2013.

2. Winkelmann R. Econometric analysis of count data. 5th ed. Berlin, Germany: Springer; 2008.

3. Manning WG, Mullahy J. Estimating log models: To transform or not to transform. Journal of Health Economics. 2001;20(4):461–94.

4. Santos Silva JMC, Tenreyro S. The log of gravity. Review of Economics and Statistics. 2006;88(4):641–58.

5. Angrist JD, Pischke J-S. Mostly harmless econometrics: An empiricist's companion. Princeton, NJ: Princeton University Press; 2009.

6. Gourieroux C, Monfort A, Trognon A. Psuedo maximum likelihood methods: Applications to Poisson models. Econometrica. 1984;52(3):701–20.

7. Barkowski S. Interpretation of nonlinear difference-in-differences: the role of the parallel trends assumption. June 2021. Available at SSRN: doi.org/10.2139/ssrn.3772458.

8. Oehlert GW. A note on the delta method. American Statistician. 1992;46(1):27–9.

Appendix B: Model validation and sensitivity analyses

Model validation

Estimates from the triple-difference design are valid under a variant of the standard difference-in-differences parallel trend assumption. Where the difference-in-differences approach requires parallel trends in the outcome of interest in the treated and untreated groups, in the absence of treatment, the triple difference design requires parallel trends only in the *difference* between the affected and unaffected outcomes in the treated and untreated groups [1-3]. For our setup, this equates to parallel trends in the (log) difference of (affected) LBP claims and of (unaffected) non-musculoskeletal injury claims in the FREE-trained (treated) cohort compared to the national population (untreated) cohort: any determinants of injury insurance claims other than FREE training are assumed not to have had differential impacts on LBP relative to non-musculoskeletal injuries for the trained cohort compared to the non-trained cohort. Note that the assumed multiplicative structure of the data (see Appendix A) is therefore reflected in the parallel trends assumption (we assess parallel trends in the log scale, and interpret the treatment effects accordingly) [4].

To assess this assumption, we estimated event study models based on the equation:

$$E[y_{ijkt}] = \exp\left(\gamma_{kt} + \lambda_{jt} + \theta_{jk} + \sum \delta_q Q_{itq} Z_{jk}\right),$$

where Q_{itq} is a dummy indicating quarter q relative to training for individual i in period t ($Q_{itq} = 0 \ \forall \ q$ for non-trained GPs) and Z_{jk} is a dummy that takes value 1 for FREE-trained GPs' LBP claims outcomes, 0 otherwise. These estimates are aggregated to the quarterly level to reduce noise in the month-by-month data. The parameters of interest are the coefficients δ_q on the quarter-by-subgroup interaction terms; as shown in Figures B1 & B2, there were no clear differences in pre-training trends for any outcomes, providing support for our difference-in-difference-in-differences identification strategy.

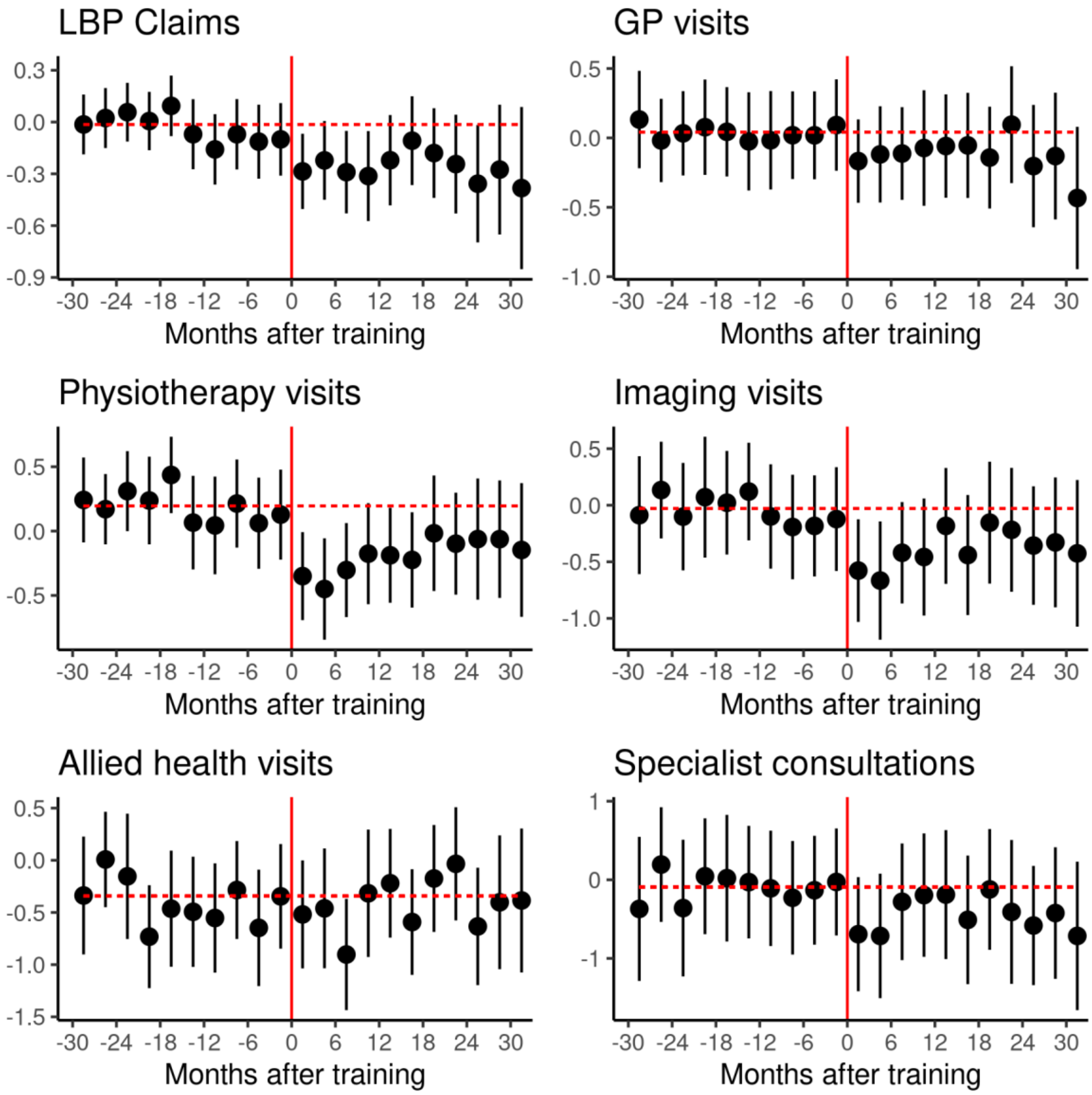

Figure B1: Timing of effect estimates, LBP claims numbers and healthcare utilisation

Each point represents a quarter-specific point estimate from the event study model. Error bars represent 95% confidence intervals. The horizontal dashed line indicates the pre-training period mean value. The first pre-training period (not shown) is the reference category.

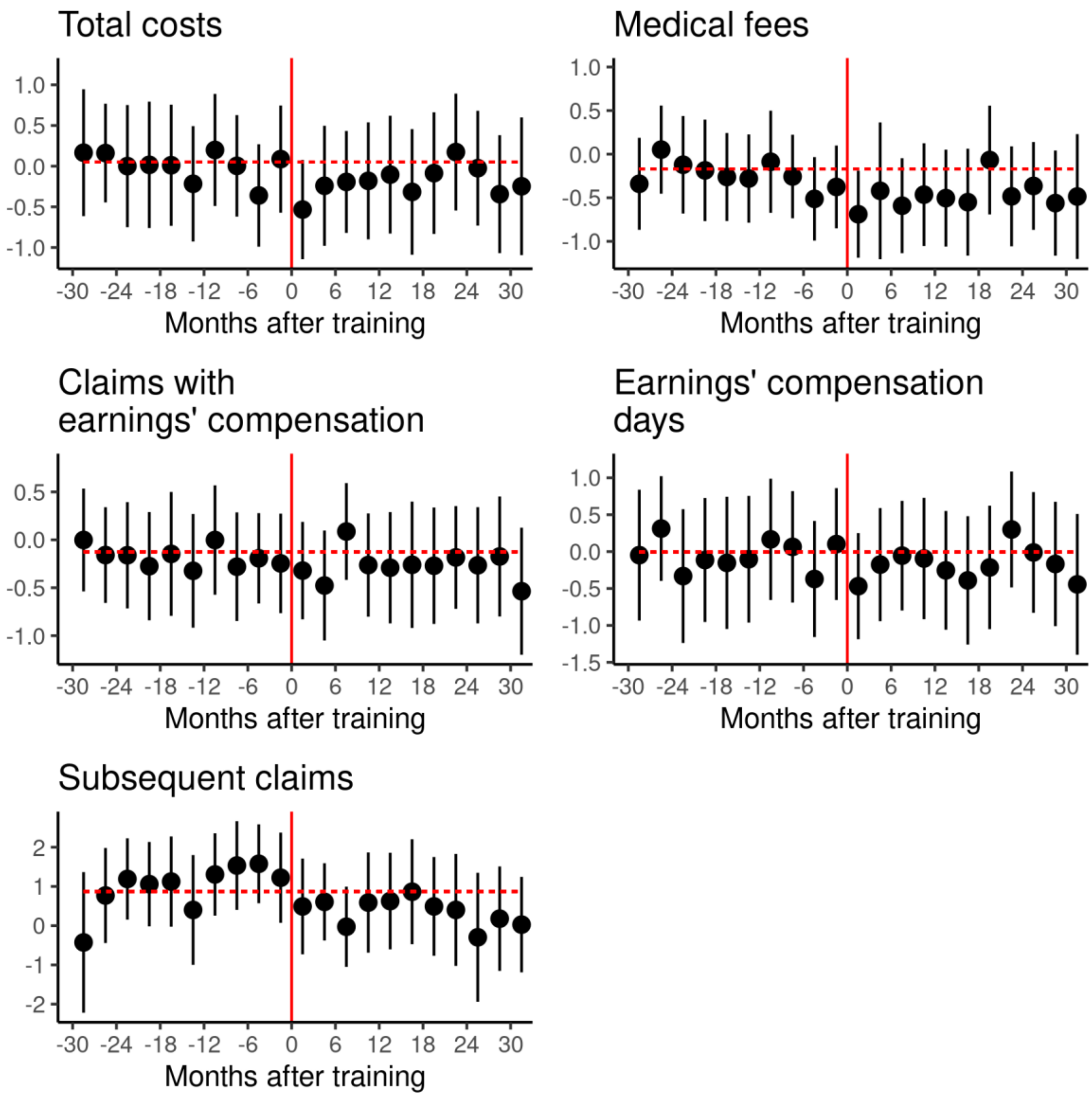

Figure B2: Timing of effect estimates, costs, earnings' compensation, and subsequent claims

Each dot represents a quarter-specific point estimate from the event study model. Bands represent 95% confidence intervals. The horizontal dashed line indicates the pre-training period mean value. The first pre-training period (not shown) is the reference category.

Another potential threat to the validity of the estimation strategy is potential self-selection of patients to GPs on the basis of FREE training. Patients are free to choose their healthcare provider; if they have strong preferences for imaging, interventions, or referrals discouraged by the FREE approach, they may move to another provider more willing to provide these services. If this behaviour is common, our estimates will overestimate the system-level impact on healthcare costs and utilisation. Evidence from the FREE RCT and implementation study suggests a mixed reception from patients to the teachings of the FREE approach: many patients felt reassured by their consultation with FREE-trained GPs and confident in following guideline-recommended LBP management strategies, while some were dissatisfied and expected a specific pathoanatomical diagnosis and stronger medical intervention. We believe this is unlikely to have substantial impact on our results for several reasons. Patients were not aware, prior to seeking care, whether their usual GP was trained in the FREE approach, so would have no reason to seek out a different provider for first consultation.[1] If an ACC claim was lodged by the FREE-trained GP at initial consultation, any subsequent care for that injury will be linked to the claim record, regardless of whether the service was referred by the original GP. Furthermore, the number of patients expressing dissatisfaction was small [5]: five percentage points higher in the FREE-trained group than the control group immediately after the initial consultation (and no significant differences between groups at subsequent follow-ups); and there was no evidence of an increase in LBP claims activity by non-FREE-trained GPs in areas with FREE training, compared with those in areas without any FREE training (Figure B3), suggesting absence of patient self-selection.

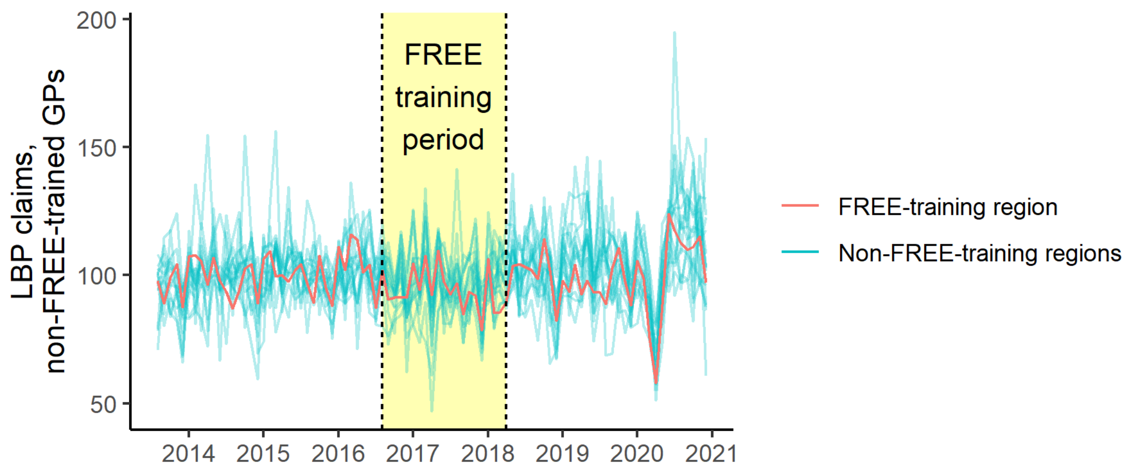

Figure B3: LBP claims by non-FREE-trained GPs, regions with and without FREE training

Values for all regions normalised to have mean value 100 in the pre-training period.

Highlighted region indicates the period (July 2016–August 2018) over which FREE training took place.

Sensitivity analyses

We conducted several robustness checks on our model specification and sample selection assumptions, summarised in Table B1 (alongside the base case analysis in Column (1) for comparison).

First, we excluded GPs practising in the Greater Wellington region from the national GP cohort, to exclude possible 'spillover' effects of training (for example, through local professional networking). These results were very similar to those of our primary analysis (Column (2)).

[1] In the NZ primary care system, patients are enrolled with practices who receive capitation payments to subsidise the patient's care. This acts as a disincentive to visit a different practice for a given consultation, as this would require paying a higher fee (the full non-subsidised cost of treatment) or changing enrolment to the new practice. The RCT was cluster-randomised at the practice level, and many GPs in the implementation study were trained within practices, so switching GPs within a practice would (in most cases) not result in seeing a non-trained GP.

Next, we re-ran the analyses using only data on claims lodged up to 18 months post-training (or prior to September 2019, 18 months after the last FREE training date, for the non-FREE-trained cohort). In March 2020 (i.e. within the 6-month follow-up period of claims lodged in September 2019), NZ entered a strict nationwide lockdown in response to the Covid-19 pandemic, resulting in a sharp discontinuity in ACC claims activity and other healthcare utilisation. Results for this sample, reported in Column (3), were very similar to our main results, with slightly stronger effects estimated for several outcomes (as would be expected given that treatment effects reduced over time in the main analysis).

Preliminary descriptive analysis identified that mean claims numbers, healthcare utilisation, and costs were notably higher in the national GP cohort than among the FREE-trained GPs, due to the existence of a long tail of GPs with extremely high claims activity. Excluding observations (GP-months) with abnormally high numbers of either LBP or non-musculoskeletal claims (more than 10 LBP or 100 non-musculoskeletal injury claims in a month) resulted in much more comparable baseline distributions for the trained and non-trained cohorts (Table B2). The point estimates of the intervention effects for this sample (Column (4)) were slightly smaller than our primary results, but were more precisely estimated; all of our main findings were unchanged.

Healthcare use, earnings' compensation, cost of claims, and subsequent claims were recorded over 6 months following the original index claim date. For claims lodged in the 6 months prior to training, these outcomes may therefore be potentially affected by training. We ran a robustness check excluding claims lodged in the last 6 months before training for all FREE-trained GPs: results were similar to or slightly stronger than in our primary analysis, consistent with potential effects of training on the outcomes of claims lodged shortly before the training date.

Aggregating the monthly data extracted from claims records to the quarterly level, to reduce the sparsity of the data for rare outcomes, made no substantive difference to any results (Column (6)). To assess the robustness of our findings to possible violation of triple-difference parallel trends, we also estimated models allowing for linear time trends (outcome-specific trends common to the trained and non-trained cohorts, and cohort-specific trends common to LBP and non-musculoskeletal claims outcomes) or group-specific linear time trends (i.e. separate time trends for each combination of cohort and outcome measure). These results (Columns (7) and (8)) were generally similar to our primary analysis. Reductions in GP visits and total claims costs became slightly larger and marginally significant (at the 10% level only) in both of the analyses with time trends, while the marginally-significant reduction in specialist consultations was no longer quite significant after accounting for group-specific trends (although the point estimate was unchanged). There were no substantive changes to results for any of the other outcomes.

Table B1: Triple-difference estimates of change in LBP claims activity following FREE training, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Base case	Excluding Wellington region	Pre-Covid lockdown period only	Excluding outliers	Excluding 6 months pre-training	Quarterly aggregation	Linear time trends	Group-specific linear time trends
Panel A								
Number of LBP claims	−0.22 (0.09) **	−0.22 (0.09) **	−0.24 (0.09) **	−0.16 (0.05) ***	−0.23 (0.08) ***	−0.23 (0.09) ***	−0.23 (0.09) **	−0.13 (0.07) *
Panel B. Number of healthcare visits								
GP	−0.14 (0.11)	−0.15 (0.11)	−0.16 (0.11)	−0.10 (0.08)	−0.13 (0.10)	−0.16 (0.11)	−0.18 (0.11) *	−0.20 (0.11) *
Physiotherapy	−0.37 (0.09) ***	−0.38 (0.09) ***	−0.50 (0.09) ***	−0.33 (0.07) ***	−0.39 (0.08) ***	−0.39 (0.09) ***	−0.40 (0.09) ***	−0.50 (0.09) ***
Imaging	−0.32 (0.13) **	−0.33 (0.13) ***	−0.43 (0.13) ***	−0.26 (0.08) ***	−0.35 (0.12) ***	−0.35 (0.13) ***	−0.36 (0.13) ***	−0.28 (0.12) **
Allied health	−0.04 (0.12)	−0.05 (0.12)	−0.16 (0.14)	0.00 (0.09)	−0.08 (0.12)	−0.06 (0.13)	−0.04 (0.13)	−0.09 (0.14)
Specialist consultations	−0.32 (0.18) *	−0.33 (0.19) *	−0.35 (0.19) *	−0.26 (0.15) *	−0.31 (0.19) *	−0.34 (0.19) *	−0.33 (0.19) *	−0.32 (0.21)
Panel C. Cost of claims								
Total	−0.21 (0.13)	−0.21 (0.13) *	−0.32 (0.14) **	−0.15 (0.10)	−0.26 (0.12) **	−0.24 (0.13) *	−0.22 (0.13) *	−0.29 (0.16) *
Medical fees	−0.26 (0.12) **	−0.26 (0.12) **	−0.36 (0.14) **	−0.21 (0.09) **	−0.33 (0.12) ***	−0.28 (0.13) **	−0.24 (0.13) *	−0.30 (0.15) **
Panel D. Earnings compensation								
Number of claims receiving earnings' compensation	−0.11 (0.13)	−0.12 (0.13)	−0.13 (0.14)	−0.04 (0.11)	−0.13 (0.12)	−0.13 (0.13)	−0.15 (0.13)	0.03 (0.15)
Total days earnings' compensation	−0.15 (0.15)	−0.16 (0.15)	−0.26 (0.16)	−0.08 (0.12)	−0.19 (0.14)	−0.19 (0.15)	−0.19 (0.15)	−0.19 (0.19)
Panel E. Subsequent claims								
Number of subsequent claims	−0.64 (0.41)	−0.64 (0.41)	−0.55 (0.42)	−0.57 (0.40)	−0.51 (0.37)	−0.66 (0.41)	−0.66 (0.41)	−1.51 (0.98)
N (trained GPs/observations)	168 / 9 073	168 / 9 073	167 / 7 153	168 / 9 043	168 / 8 203	168 / 3 219	168 / 9 073	168 / 9 073
N (non-trained GPs/observations)	6 075 / 311 142	5 450 / 279 971	6 074 / 262 346	6 030 / 301 285	6 075 / 311 142	6 075 / 116 199	6 075 / 311 142	6 075 / 311 142
FREE-trained cohort sample	All	All	Claims lodged up to 18 months post-training	Excluding outliers	Excluding claims lodged within 6 months pre-training	All	All	All
Non-trained cohort sample	All	Excluding Greater Wellington region	Claims lodged before September 2019	Excluding outliers	All	All	All	All
Aggregation	Monthly	Monthly	Monthly	Monthly	Monthly	Quarterly	Monthly	Monthly
Model	Non-parametric controls	Non-parametric controls	Non-parametric controls	Non-parametric controls	Non-parametric controls	Non-parametric controls	Linear time trends	Group-specific linear time trends

Standard errors, clustered at the GP level, are shown in parentheses. p-values: *** < 0.01 < ** < 0.05 < * < 0.1.

Table B2: Descriptive statistics of GP cohorts, excluding GP-months with abnormally high numbers of ACC claims

	FREE-trained cohort		Non-trained cohort	
Outcome	Pre-training	Post-training	Pre-training	Post-training
LBP claims				
Number of LBP claims	30.5 (25.1)	26.7 (20.4)	29.6 (36.1)	26.2 (32.8)
Number of healthcare visits on LBP claims				
GP	17.2 (19.1)	17.0 (17.1)	23.0 (38.2)	21.6 (36.1)
Physiotherapy	48.7 (46.4)	34.6 (28.6)	50.7 (68.1)	41.8 (56.7)
Imaging	6.7 (7.6)	5.7 (5.4)	7.9 (11.4)	7.6 (11.4)
Allied health	21.2 (26.4)	24.1 (29.4)	40.8 (109.3)	39.6 (107.6)
Specialist consultations	3.3 (4.4)	2.9 (5.1)	4.6 (7.8)	4.5 (7.9)
Cost of LBP claims				
Total	$19 783 ($24 888)	$21 471 ($23 087)	$23 584 ($36 880)	$24 124 ($38 365)
Medical fees	$6 169 ($7 584)	$5 736 ($6 020)	$9 230 ($14 686)	$8 718 ($13 808)
Earnings' compensation on LBP claims				
Number of claims receiving earnings' compensation	1.9 (2.4)	2.0 (2.2)	2.0 (3.1)	1.9 (3.0)
Total days earnings' compensation	98 (144)	108 (134)	112 (195)	113 (200)
Claims subsequent to LBP claims				
Number of subsequent claims	0.8 (4.8)	0.5 (1.0)	0.9 (1.8)	0.8 (1.6)
Number of non-MSK claims filed	163.9 (157.2)	155.9 (149.7)	167.6 (245.4)	141.8 (211.7)

Values are mean (SD). Data collected over 33 months before and after training date for FREE-trained GPs; 33 months before and after the median training date (November 2017) for non-trained GPs.

'Abnormally high' numbers of ACC claims are more than 10 LBP claims or 100 non-musculoskeletal claims per month.

References

1. Olden A, Møen J. The triple difference estimator. NHH Institutt for foretaksøkonomi; 2020 Apr.

2. Deschênes O, Greenstone M, Shapiro JS. Defensive investments and the demand for air quality: Evidence from the NOx Budget Program. American Economic Review. 2017;107(10):2958–89.

3. Walker WR. The transitional costs of sectoral reallocation: Evidence from the Clean Air Act and the workforce. Quarterly Journal of Economics. 2013;128(4):1787–835.

4. Barkowski S. Interpretation of nonlinear difference-in-differences: the role of the parallel trends assumption. June 2021. Available at SSRN: doi.org/10.2139/ssrn.3772458.

5. Darlow B, Stanley J, Dean S, Abbott JH, Garrett S, Wilson R, et al. The Fear Reduction Exercised Early (FREE) approach to management of low back pain in general practice: A pragmatic cluster-randomised controlled trial. PLoS Medicine. 2019;16(9):e1002897.

Appendix C: Supplementary data and results

Tables

Table C1: LBP injury Read Codes used for data extraction

Code	Description
S5610	Sprain, anterior sacro-iliac ligament
SK191	Traumatic rupture of lumbar intervertebral disc
N1466	Sacroiliac disorder
N23YE	Spasm of back muscles
S562.	Sprain, sacrospinous ligament
S56Z.	Sacroiliac spain NOS
SYU36	Sprain & strain of oth & unsp parts of lumbar spine & pelv
N142.	Pain in lumbar spine
N1420	Lumbago with sciatica
N145.	Backache, unspecified
S57Z0	Pulled back muscle
N122.	Lumbar disc displacement
N12C.	Disc prolapse with radiculopathy
N12C2	Lumbar disc prolapse with radiculopathy
N12Z3	Other lumbar disc disorders
N143.	Sciatica
N1463	Lumbrosacral instability
S560.	Sprain, lumbrosacral ligament
S561.	Sacroiliac ligament sprain
S57..	Sprain of other parts of back
S572.	Lumbar sprain
S573.	Sacrum sprain
S57X.	Sprain & strain of oth & unsp parts of lumb spine & pelv
S57Z.	Back sprain NOS
SJ32.	Lumbar nerve root injury

Non-musculoskeletal and non-primary care managed injuries included burns, fractures/dislocations, hernias, lacerations/wounds, hearing loss, poisonings, dental injuries, concussions, medical treatment injuries, gradual onset conditions due to occupational exposure, and others (full list of included codes available on request from the authors).

Table C2: Implementation costs

Item	Cost per workshop + refresher	Cost per GP[1]
Administration	\$470.00	\$39.17
Workshop facilitator	\$944.00	\$78.67
GP training remuneration		\$800.00
GP practice consultation compensation		\$60.00
GP training manual and stand		\$15.45
Catering		\$25.00
Facilitator travel	\$85.00	\$7.08
Total		**\$1 025.37**

[1] Per-GP share of workshop costs calculated based on an average of 12 GPs per workshop.

Figures

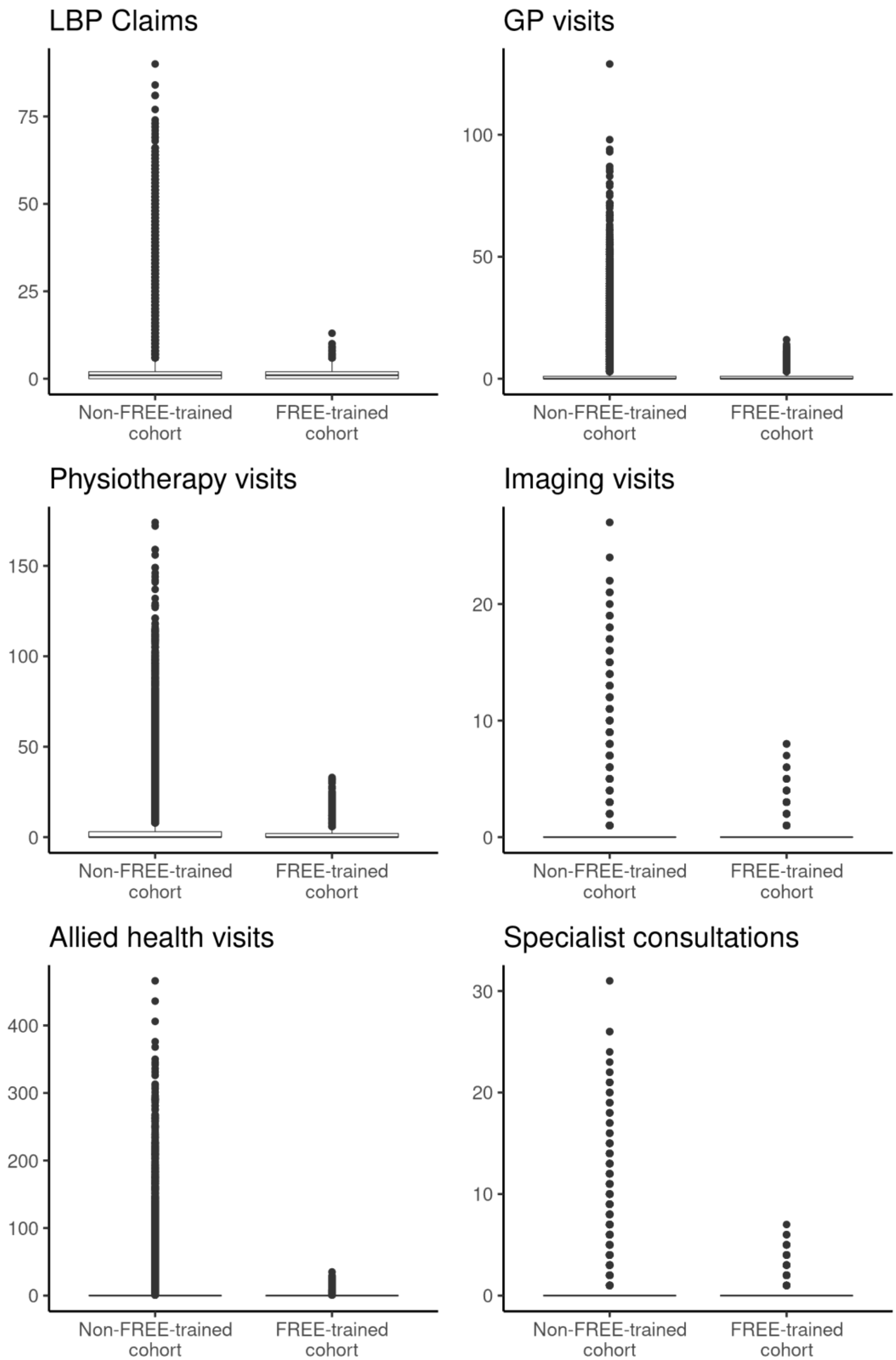

Figure C1: Baseline outcome distributions, FREE-trained and non-FREE-trained cohorts, LBP claims numbers and healthcare utilisation

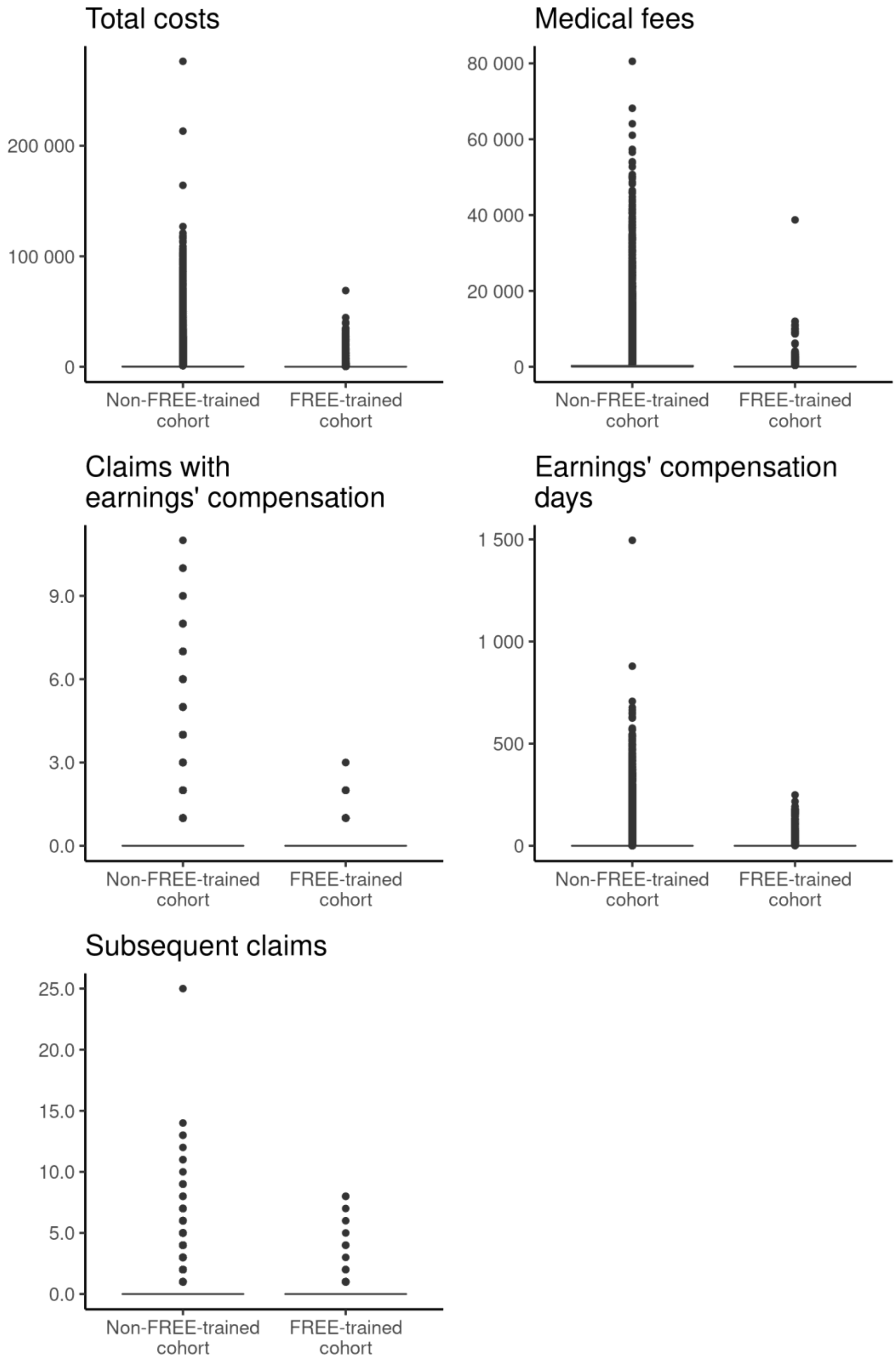

Figure C2: Baseline outcome distributions, FREE-trained and non-FREE-trained cohorts, LBP claims costs, earnings' compensation, and subsequent claims

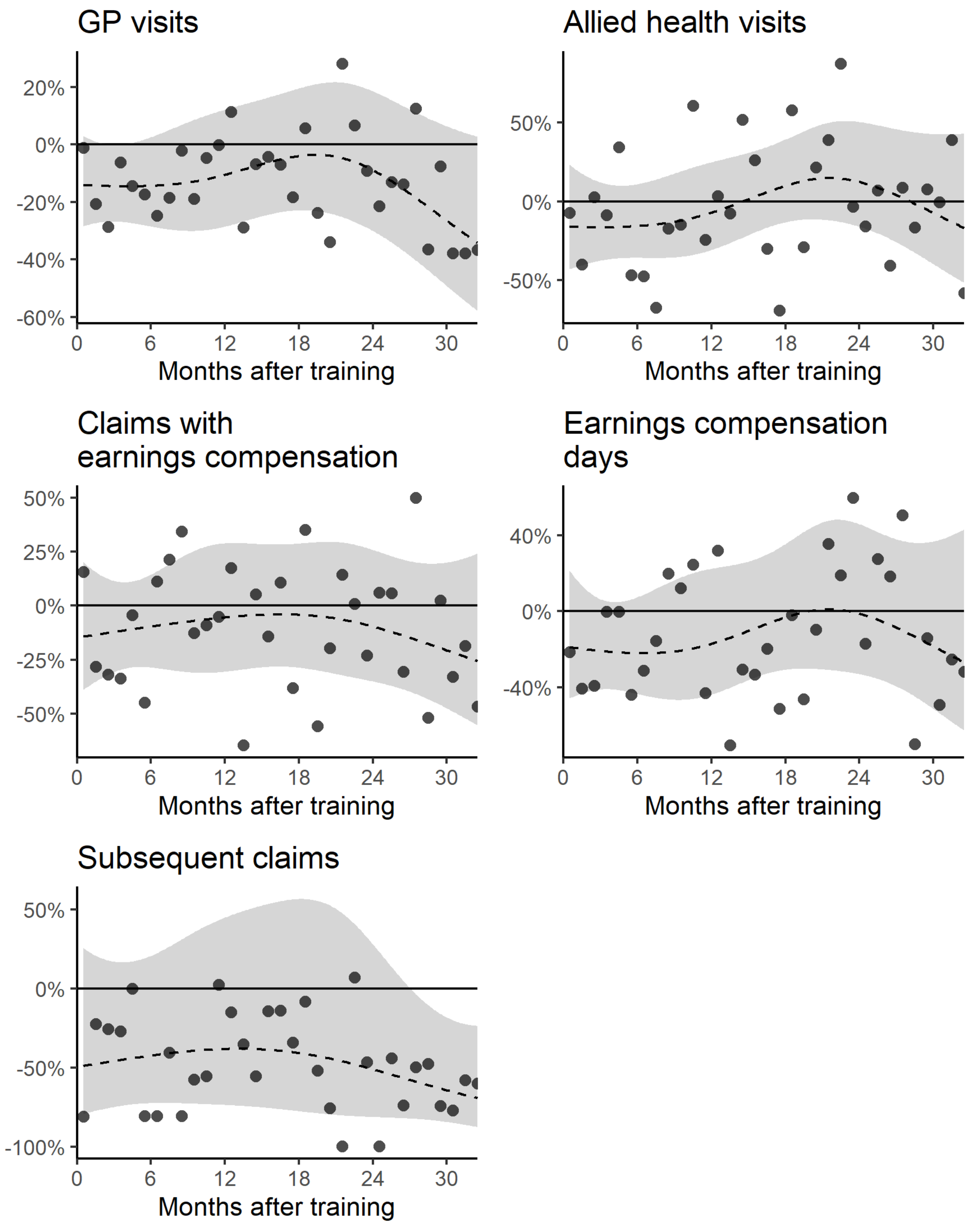

Figure C3: Trajectory of triple difference estimates of effect of FREE training (additional outcomes)

Each dot represents the month-specific treatment effect estimate. Smoothed curve and confidence band (representing 95% confidence intervals) from cubic spline smoothing function.