Cureus
Part of SPRINGER NATURE

Open Access Original Article

Harmonized Work Health and Safety Laws and Traumatic Occupational Digit Amputation in Australia: A Staggered Difference-in-Differences Study

Michael J. Franco [1]

1. Plastic and Reconstructive Surgery, Cooper University Hospital, Camden, USA

Corresponding author: Michael J. Franco, franco-michael@cooperhealth.edu

Review began 10/02/2026
Review ended 10/24/2026
Published 11/01/2026

© **Copyright** 2026
Franco. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118943

Abstract

Objective

This study aimed to determine whether Australia's staggered Work Health and Safety (WHS) harmonization was associated with reduced accepted workers' compensation claims for traumatic occupational digit (finger and thumb) amputation.

Methods

This near-national, administrative claims-based difference-in-differences study covered eight Australian workers' compensation jurisdictions. The primary analysis compared the 2012 and 2013 WHS adoption cohorts with Victoria (never treated) and Western Australia (WA; not yet treated) as comparators. In the pooled model, jurisdictions that had already adopted the laws also served as controls for later adopters. A prespecified exploratory analysis separately evaluated WA's 2022 adoption against Victoria. Claim counts were modeled using Poisson models with jurisdiction and financial-year fixed effects and a log(jobs) offset. Because only six to eight jurisdictions were available, inference relied on exact randomization tests (RI) rather than asymptotic cluster-robust standard errors.

Results

The analytic sample contained 5,259 primary-outcome claims across 73 jurisdiction-years (the study window recorded 5,502 claims, of which 243 fell within mixed-exposure years excluded by design). The pooled incidence-rate ratio (IRR) was 1.03 (descriptive 95% CI: 0.85-1.25) overall and 1.00 (0.83-1.20) when excluding New South Wales (NSW). Under RI, neither estimate showed evidence of an association ($p = 0.89$ and $p = 0.98$, respectively). A naïve pretrend test was significant ($p = 0.012$), but the corresponding RI test did not detect differential pretrends ($p = 0.48$). Results were likewise null for machinery-related digit amputation (IRR: 0.99, $p = 0.96$) and the broader time-loss outcome (IRR: 0.99, $p = 0.80$). The overall interpretation remained unchanged across sensitivity analyses, including a stacked specification that prevented already-treated jurisdictions from serving as controls.

Conclusions

I found no evidence that harmonizing WHS laws was associated with reduced accepted workers' compensation claims for digit amputation among the primary adoption cohorts, but a simulation showed 80% power only for reductions of roughly 40% to 55% or larger. One possible interpretation is that harmonization occurred against an already-declining and already regulated hazard, leaving limited additional room for measurable reductions attributable to the reform.

Categories: Health Policy, Orthopedics, Epidemiology/Public Health
Keywords: australia, difference-in-differences, digit amputation, natural experiment, occupational injury, randomization inference, workers' compensation, workplace regulation

Introduction

Traumatic amputation of the fingers or thumb is a common, disabling, and mechanistically well-understood form of occupational hand trauma. Global Burden of Disease (GBD) modeling estimated 1.86 million incident thumb amputations and 4.30 million incident non-thumb digit amputations worldwide in 2017 alone [1]. More recent GBD analyses identified finger (non-thumb) amputation as the most prevalent amputation subtype globally, accounting for 50.3% of all cases in 2021 [2]. In GBD 2017, Australasia (the region comprising Australia and New Zealand) had the highest regional age-standardized incidence of both thumb amputation (74.2 per 100,000) and non-thumb digit amputation (207.5 per 100,000) among all GBD regions worldwide [1]. Australia's substantial occupational injury burden, national workers' compensation infrastructure, and staggered implementation of Work Health and Safety (WHS) harmonization together

make it a useful setting for evaluating whether this policy change was associated with a change in severe occupational injury.

Most occupational digit amputations result from mechanical force, particularly contact with machinery, which makes them potentially responsive to engineering and administrative machine-safety controls. In a 20-year descriptive analysis of Victorian hospital admissions, mechanical forces caused 95% of work-related digit amputations in the most recent decade [3], a pattern that has also been reported internationally [4,5].

Before harmonization, Australian workplace health and safety law was set independently by each of the country's eight states and territories. In 2008, the Council of Australian Governments signed an intergovernmental agreement to harmonize this framework. Safe Work Australia then drafted a model WHS Act, which the Workplace Relations Ministers' Council endorsed in December 2009, with jurisdictions expected to enact it by January 1, 2012 [6]. New South Wales (NSW), Queensland (QLD), the Australian Capital Territory (ACT), and the Northern Territory (NT) implemented the model laws on January 1, 2012; South Australia (SA) and Tasmania (TAS) followed on January 1, 2013; and Western Australia (WA) adopted a modified version on March 31, 2022. Victoria has not adopted the model WHS Act. This staggered rollout creates quasi-experimental variation that can be used to evaluate the association between harmonization and occupational injury incidence.

Safe Work Australia's 2011 evaluation plan identified change in the rate of work-related serious injury and death as an outcome measure for the harmonization process [6]. However, to my knowledge, no published study has directly evaluated harmonization against a severe injury endpoint using a quasi-experimental design. Existing evaluations are either perception-based [7,8] or address adjacent outcomes. A stacked difference-in-differences analysis of self-reported survey data found no significant pooled effect on workers' compensation receipt overall in its main specification (which excluded NSW), although a reduction was detected in high-risk industries [9]. An interrupted time-series analysis of the same national dataset found that a NSW benefit-scheme reform in June 2012 (the year NSW harmonized its WHS laws) reduced accepted NSW claims by 15.3% relative to seven comparator jurisdictions [10]. A descriptive analysis documented a substantial decline in Victorian digit amputation incidence over 20 years [3]. Because Victoria never adopted the harmonized laws, the decline there illustrates why a cross-jurisdiction design, rather than a simple before-and-after comparison, is needed to isolate an association with harmonization.

To clarify the policy exposure, the model WHS Act imposed a duty on a person conducting a business or undertaking to ensure, so far as reasonably practicable, the health and safety of workers and other persons put at risk by work, and it broadened the definitions of the employment relationship and the workplace [9]. Other changes included higher maximum financial penalties (the maximum jurisdiction-level penalty was doubled), the introduction of criminal penalties, and written enforceable undertakings that a regulator could accept in lieu of prosecution. In NSW and Queensland, the reverse onus of proof was also removed [9]. Implementation and enforcement nevertheless remained with each jurisdiction's regulator, and the model Act allowed jurisdictions to customize aspects of administration and enforcement [9]. Exposure to the reform therefore differed across jurisdictions in timing (2012, 2013, or 2022), in content (WA adopted a modified version), and in whether it occurred at all (Victoria).

The primary objective was to estimate whether adoption of the harmonized WHS laws was associated with a change in the incidence of accepted workers' compensation claims for traumatic digit amputation among the 2012 and 2013 adoption cohorts. Secondary objectives were to test whether any association was concentrated in machinery-related mechanisms; to evaluate WA's own 2022 adoption separately against Victoria (a prespecified exploratory analysis); and to address a small-cluster inference problem relevant to similarly scaled designs. Because jurisdictions adopted on different dates, a staggered difference-in-differences design compares the change in each adopting jurisdiction with the change in jurisdictions that had either not yet adopted or never adopted, thereby removing stable between-jurisdiction differences and shared national time trends.

The primary analysis compares the 2012 and 2013 adoption cohorts against Victoria (never treated) and WA (not yet treated within the primary window). In the pooled model, already-adopting jurisdictions also serve as controls for later adopters. The design rests on three assumptions: parallel trends (absent harmonization, adopting and comparator jurisdictions would have followed similar trends), no anticipation (claims did not change before adoption in response to it), and no important concurrent jurisdiction-specific interventions. None can be verified directly, but the Methods and Limitations sections describe how each was assessed.

Materials And Methods

I used administrative workers' compensation claims data from Safe Work Australia's National Dataset for Compensation-based Statistics (NDS), obtained via a custom data request (Submission 875) covering financial years (FY) 2009-10 through 2023-24 (provisional). The NDS aggregates claims across Australia's eight states and territories; Comcare and Seacare (the Commonwealth schemes for federal employees and maritime workers) are not geographically bounded and were excluded, so this dataset covers near-national workers' compensation claims across the eight jurisdictions, not the entire Australian labor force. The Australian Capital Territory reports separate private- and government-sector schemes; the extract provided

a claims numerator only for the ACT private-sector scheme, so the ACT unit here refers to that scheme, matched to the corresponding private-sector jobs denominator.

The offset denominator ("jobs") was supplied directly by Safe Work Australia within the same NDS extract as the claims numerator, matched to each jurisdiction-financial-year cell; for ACT it is restricted to the private-sector series to match the private-sector claims numerator. Separate Australian Bureau of Statistics (ABS) Labour Force data were used only to construct the industry-composition, unemployment-rate, and female-employment-share covariates used in the robustness analysis below, not the primary offset.

Each observation in the analysis is therefore one jurisdiction-financial-year cell containing the number of accepted workers' compensation claims meeting the outcome definition and the matching jobs denominator. Claims in the NDS are classified using the national Type of Occurrence Classification System (TOOCS), which supports comparable outcome coding across jurisdictions. Compensation coverage, claim eligibility rules, and acceptance practices nonetheless differ between schemes and can change over time, as the 2012 NSW benefit-scheme reform illustrates [10].

Because every model includes jurisdiction fixed effects, differences between jurisdictions that are constant over time (for example, scheme coverage, eligibility thresholds, coding level, and how the denominator is defined) are absorbed and do not bias the estimates. Changes within a jurisdiction over time are not absorbed and are a threat to validity (Limitations); I did not independently verify the comparability of the denominator across jurisdictions and years beyond using the series supplied in the extract. The ACT private-sector restriction narrows that unit's coverage and, because ACT contributes few unsuppressed cells, a sensitivity analysis omitting NT and ACT gave a similar estimate.

Study design and treatment cohorts

I used a staggered difference-in-differences design exploiting the jurisdiction-by-jurisdiction rollout of Australia's harmonized WHS model laws. NSW, QLD, ACT and NT implemented the harmonized laws on 1 January 2012 ("2012 cohort"); SA and TAS implemented on 1 January 2013 ("2013 cohort") [9]. These two cohorts, evaluated against Victoria (never treated) and WA (not yet treated within the primary window), form the primary analysis. WA implemented a modified version on 31 March 2022 (sources for all adoption dates: Supplementary File, Section A; see Google Doc link in the Appendices); since it cannot contribute treated observations to the primary window, I evaluate it separately in a prespecified exploratory analysis against Victoria alone.

The design rests on three assumptions, none of which can be verified directly. Parallel trends (absent harmonization, adopting and comparator jurisdictions would have followed similar trends) were assessed indirectly through pretrend diagnostics and event-study estimates. No anticipation (claims did not change before adoption in response to it) was examined using the pre-adoption event-study coefficients. The absence of important concurrent jurisdiction-specific interventions was addressed through the exclusion of NSW, which had a contemporaneous benefit-scheme reform, and through leave-one-out analyses (Limitations).

I coded exposure as binary, beginning in the first full financial year after each jurisdiction's adoption date, and excluded the financial year containing the adoption date as a mixed-exposure period. This estimates the association with statutory adoption (an intention-to-treat contrast). It does not capture differences between jurisdictions in implementation intensity, regulator enforcement, employer compliance, or delayed effects, which could not be measured with the available data. The event-study specifications relax the binary assumption by estimating a separate coefficient for each year relative to adoption, so that delayed or growing effects could appear; they showed no sustained decline (Results).

Outcome definitions

The primary outcome was traumatic finger or thumb amputation, defined as accepted claims coded to TOOCS Nature of injury 139 and Bodily location 461/462, by date of occurrence. The machinery-related digit-amputation outcome (protocol outcome C) further required TOOCS Agency major group 1 (machinery) or Agency 3119 (powered hand tools). The broader machinery-related time-loss hand-injury outcome (protocol outcome E) included accepted claims with at least one week of compensated time lost and Bodily location 450, 460-462, 468, or 469, with the same machinery/powered-tool restriction. Permanently disabling machinery-related hand injury (protocol outcome F) could not be estimated at the jurisdiction level because the Safe Work Australia extract provided this outcome only pooled across jurisdictions. Because all ACT machinery-related cells were suppressed, machinery-related analyses were estimated in seven jurisdictions.

Panel construction

I constructed a panel with one row per jurisdiction-financial-year. The primary analysis window was restricted to FY2009-10 through FY2019-20, limiting the contribution of the major workplace disruptions of the coronavirus disease 2019 (COVID-19) era in subsequent financial years while preserving Victoria and WA as valid untreated comparators throughout. Safe Work Australia suppresses jurisdiction-year cells with

fewer than five claims (NP); these were treated as missing and excluded row-wise, bracketed by a sensitivity analysis replacing NP with 0 and separately with 4. Each jurisdiction's implementation date fell mid-financial-year; that financial year was excluded as a mixed-exposure period, leaving only two to three fully pre-adoption years for the 2012/2013 cohorts (Discussion).

Statistical analysis

Jurisdiction-by-year claim counts were modeled with a Poisson generalized linear model with jurisdiction and financial-year fixed effects and a log(jobs) offset (Equation 1):

$$\log \mathrm{E}[Y_{jt}] = \alpha_j + \lambda_t + \beta \cdot \mathrm{Post}_{jt} + \log(\mathrm{Jobs}_{jt})$$

where Y is the claim count for jurisdiction j in financial year t, α and λ are jurisdiction and financial-year fixed effects, Post is 1 for fully post-adoption jurisdiction-years and 0 otherwise, and Jobs is the jobs denominator, entered as an offset with its coefficient fixed at 1. The parameter of interest is β, and exp(β) is the incidence-rate ratio (IRR). Event-study specifications replaced Post with cohort-specific relative-time indicators (reference period: two years before adoption) to examine dynamic associations and pre-adoption trends, motivated by concerns about treatment-effect heterogeneity in staggered-adoption designs.

The original protocol specified adjustment for time-varying jurisdiction-level industry composition, unemployment, and sex composition; the covariate-adjusted models included the manufacturing, construction, and agriculture shares of employment, the unemployment rate, and the female share of employment (ABS Labour Force data). The bare model (jurisdiction and financial-year fixed effects without covariates) was used as the primary specification because adjustment for variables potentially influenced by, or evolving concurrently with, treatment can introduce bias. The covariate-adjusted specification was retained as a robustness analysis and changed the primary estimate only modestly (IRR 1.03 to 1.08; Results).

In this static two-way fixed-effects model, the post-adoption coefficient is a weighted average of comparisons between adopting jurisdictions and Victoria and WA (never or not yet treated) and also of comparisons in which already-adopting jurisdictions serve as controls for later adopters [11]. To show whether the latter comparisons drive the result, I report a clean-control stacked specification in which each cohort is compared only with Victoria and WA. The two estimates agree (IRR 1.03 versus 1.04). I kept the static model as primary because it uses all jurisdiction-years in a single summary estimate, and I report the stacked estimate alongside it in the Abstract, Results, and Discussion.

With only six to eight clusters, asymptotic cluster-robust standard errors are unreliable [12,13]; naïve tests returned implausibly small pretrend p-values (e.g., $p < 10^{-20}$). I therefore based inference on exact randomization tests (randomization inference, RI), enumerating every relabeling of which jurisdictions received which treatment-timing pattern (420 relabelings for the full eight-jurisdiction specification) and computing p-values as the share of relabelings at least as extreme as observed. "Exact" describes the enumeration, not literal random policy assignment: the test asks how unusual the observed pattern is against the null that treatment-timing labels are exchangeable across jurisdictions, conditional on true group sizes (formal statement in Supplementary File, Section A; see Google Doc link in the Appendices). The randomization tests are valid if, conditional on the observed group sizes (four, two, and two jurisdictions) and after accounting for jurisdiction and year effects, the assignment of adoption-timing patterns to jurisdictions is exchangeable under the null hypothesis of no association.

This is an assumption, not a feature of the design: adoption dates were set by each legislature, not randomly, and the assumption would fail if earlier-adopting jurisdictions had systematically different underlying injury trends. Three considerations support it here, although none proves it. First, the intergovernmental agreement set a common target date of 1 January 2012 [6], and I am not aware of evidence that the order of adoption was related to jurisdictions' injury trends. Second, the randomization-based pretrend test, which is sensitive to differential pre-adoption trends between early and late adopters, did not detect one (Results). Third, the conclusions were the same when the largest jurisdiction (NSW) was excluded (210 relabelings) and under the t-statistic variant, which is less sensitive to heterogeneous cluster sizes [14]. With only two to three pre-adoption years, the pretrend test has limited power, and the p-values should be read as conditional on the exchangeability assumption.

This coefficient-based procedure (RI-β) can be poorly calibrated when clusters differ substantially in size [14], so I also implemented a t-statistic-based variant (RI-t) for the primary, machinery-related, broader time-loss, and triple-difference outcomes. Naïve cluster-robust confidence intervals (CIs) are descriptive only; conclusions rest on the randomization-inference p-values. For the machinery-related outcomes, in which all ACT cells were suppressed, the same conclusions held when randomization inference was restricted to the seven contributing jurisdictions (210 relabelings; Supplementary File, Section B5; see Google Doc link in the Appendices).

Analyses specified in the protocol before data analysis were the primary Poisson model and cohorts,

exclusion of the mixed-exposure year, covariate adjustment, the machinery-related and broader time-loss secondary outcomes, the exploratory WA analysis, exclusion of NSW because of its contemporaneous benefit-scheme reform, and restriction to six jurisdictions with negligible suppression. The remaining analyses were added after initial results and are post hoc: alternative handling of suppressed cells (replacing NP with 0 and with 4); follow-up truncated at FY2015-16 and FY2018-19; leave-one-out re-estimation (bare and covariate-adjusted); the machinery triple-difference with mechanism-by-year fixed effects, which absorbs any nationwide, calendar-time-varying difference between machinery and non-machinery claim rates; the cohort-specific stacked analysis comparing each treated cohort only with never-treated or not-yet-treated jurisdictions; randomization inference, including its t-statistic variant (RI-t), which was added because cluster-robust inference is unreliable with so few clusters; and a simulation-based power analysis (below).

Follow-up through FY2018-19 ends before the COVID-19 onset that falls within FY2019-20, and truncation at FY2015-16 tested whether the estimate depended on the length of post-adoption follow-up. The stacked analysis directly addressed the general concern that staggered-adoption fixed-effects designs may use already-treated units as comparators for later-treated units [11]. A formal Goodman-Bacon decomposition was not relied on for the primary Poisson specification; the attempt and its limitations are documented in Supplementary File, Section A (see Google Doc link in the Appendices). Sensitivity analyses were fitted in the bare model and, where indicated, in the covariate-adjusted model; naïve triple-difference intervals are from the bare model with mechanism-by-year fixed effects.

Because the number of jurisdictions is small, I estimated statistical power post hoc by simulation. Starting from the fitted primary model with the post-adoption effect removed, I simulated claim counts for the observed panel structure, applying a true post-adoption IRR between 1.0 and 0.4 to the jurisdiction-years that were truly post-adoption. Counts were simulated as Poisson and, because the primary model's Pearson dispersion was 2.0, as negative binomial with that dispersion (300 simulations per scenario). Each simulated dataset was analyzed with the primary model and exact RI-β (420 relabelings), and power was the proportion of simulations with $p < 0.05$. RI-t was not simulated (Supplementary File, Section B7; see Google Doc link in the Appendices).

All analyses were conducted in Python 3.12; code and data-extraction workbooks form a reproducible pipeline. Use of AI tools is disclosed in the Acknowledgments and detailed in Section C of the Supplementary File (see Google Doc link in the Appendices).

Results

Descriptive characteristics

The analytic panel comprised eight jurisdictions over 15 financial years (FY2009-10 to FY2023-24). The primary analysis window (FY2009-10 to FY2019-20) contained 88 jurisdiction-years. Ten were suppressed by Safe Work Australia (fewer than five claims; ACT seven, NT three), and the remaining 78 reported 5,502 primary-outcome claims (Table *1*). Six jurisdiction-years were mixed-exposure years excluded by design; five of them reported counts (243 claims) and one (ACT, FY2011-12) was already suppressed. The analytic sample therefore comprised 73 jurisdiction-years and 5,259 claims (Table *2*). Crude rates ranged from 1.6 per 100,000 jobs in NSW to 8.6 in NT (Table *2*). They were lower after adoption than before in four of the six treated jurisdictions (NSW, QLD, SA, TAS) and higher in ACT and NT, which contribute few unsuppressed observations. Victoria, which never adopted the model laws, also declined, from 7.0 per 100,000 jobs in FY2009-10 to 5.6 in FY2019-20 (Table *1*). Figure *1* shows the raw rates; no jurisdiction showed an obvious, isolated pre-treatment trend break.

Financial year	NSW	QLD	ACT	NT	SA	TAS	VIC	WA	All
2009-10	1.9 (60)	6.5 (123)	4.5 (5)	7.1 (8)	6.3 (45)	5.8 (12)	7.0 (177)	9.9 (106)	5.5 (536)
2010-11	1.7 (53)	10.1 (192)	NP	8.7 (10)	8.8 (63)	8.1 (17)	7.8 (202)	9.3 (102)	6.6 (639)
2011-12	1.3 (40)†	6.1 (119)†	NP	10.3 (12)†	7.4 (53)	5.7 (12)	5.8 (149)	8.0 (92)	4.8 (477)
2012-13	1.2 (40)	6.4 (124)	5.7 (7)	9.9 (12)	8.8 (65)†	3.3 (7)†	6.8 (177)	7.2 (88)	5.1 (520)
2013-14	1.4 (45)	5.5 (117)	3.9 (5)	10.8 (14)	7.2 (52)	5.8 (12)	5.6 (145)	6.1 (75)	4.5 (465)
2014-15	1.4 (46)	5.1 (110)	NP	5.9 (8)	6.2 (45)	8.7 (19)	6.1 (165)	5.3 (66)	4.4 (459)
2015-16	1.3 (46)	4.5 (98)	NP	8.1 (11)	6.4 (47)	4.1 (9)	5.4 (149)	6.4 (78)	4.1 (438)
2016-17	1.1 (37)	6.3 (138)	4.2 (6)	9.9 (14)	6.7 (50)	5.0 (11)	5.2 (149)	7.0 (85)	4.5 (490)
2017-18	1.2 (44)	4.6 (106)	NP	NP	5.1 (39)	2.6 (6)	5.9 (175)	5.7 (70)	4.0 (440)
2018-19	2.2 (81)	5.8 (133)	NP	NP	5.5 (43)	5.7 (13)	4.9 (153)	5.1 (64)	4.3 (487)
2019-20	2.7 (97)	6.4 (148)	NP	NP	5.2 (40)	4.6 (11)	5.6 (176)	6.2 (79)	4.8 (551)
Total, 11 years	1.6 (589)	6.1 (1,408)	4.6 (23)	8.8 (89)	6.7 (542)	5.4 (129)	6.0 (1,817)	6.9 (905)	4.7 (5,502)

TABLE 1: Primary-outcome claims and crude incidence rates by financial year and jurisdiction, FY2009-10 to FY2019-20

†Mixed-exposure financial year, in which the jurisdiction's WHS adoption date fell; excluded from the analytic sample

Cells show the crude rate of accepted primary-outcome claims per 100,000 jobs, with the number of claims in parentheses. NP: not published by Safe Work Australia (fewer than five claims). All: pooled rate and claims across the jurisdictions with reported counts in that year. Total: sum of reported cells over the 11 years (78 of 88 jurisdiction-years). Rates use the Safe Work Australia jobs denominators; mean annual jobs by jurisdiction are in Table 2

NSW: New South Wales; QLD: Queensland; ACT: Australian Capital Territory; NT: Northern Territory; SA: South Australia; TAS: Tasmania; VIC: Victoria; WA: Western Australia; WHS: Work Health and Safety

Jurisdiction	WHS adoption	Years in analytic sample (n)	Claims (n)	Mean annual jobs (n)	Crude rate per 100,000 jobs	Rate before adoption	Rate after adoption
NSW	Jan 1, 2012	10	549	3,386,097	1.6	1.8	1.6
QLD	Jan 1, 2012	10	1,289	2,122,366	6.1	8.3	5.6
ACT	Jan 1, 2012	4	23	126,166	4.6	4.5	4.6
NT	Jan 1, 2012	7	77	127,328	8.6	7.9	8.9
SA	Jan 1, 2013	10	477	741,296	6.4	7.5	6.0
TAS	Jan 1, 2013	10	122	218,486	5.6	6.5	5.2
VIC	Not adopted	11	1,817	2,769,765	6.0	6.0	–
WA	Mar 31, 2022*	11	905	1,200,734	6.9	6.9	–
All jurisdictions		73	5,259	–	4.8	–	–

TABLE 2: Analytic sample by jurisdiction: claims, jobs denominators, and crude incidence rates

*WA adopted a modified version on 31 March 2022, after the primary window.

The analytic sample is the primary analysis window (FY2009-10 to FY2019-20) after exclusion of each jurisdiction's mixed-exposure financial year and of cells suppressed by Safe Work Australia (fewer than five claims). Crude rate is claims divided by jobs, summed over the included years, per 100,000 jobs. Rate before and after adoption refer to years before and after the mixed-exposure year; VIC and WA were untreated throughout the window. The 5,259 claims exclude 243 claims in five mixed-exposure years and any claims in ten suppressed cells; the window recorded 5,502 claims in 78 cells with reported counts (Table 1)

NSW: New South Wales; QLD: Queensland; ACT: Australian Capital Territory; NT: Northern Territory; SA: South Australia; TAS: Tasmania; VIC: Victoria; WA: Western Australia; WHS: Work Health and Safety

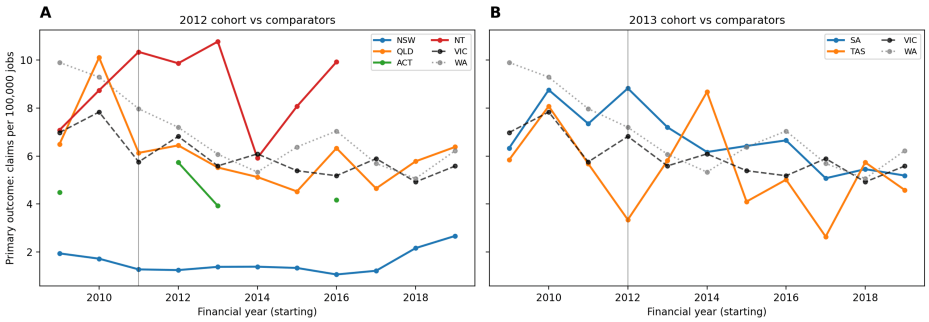


FIGURE 1: Digit amputation claim rates by jurisdiction and adoption cohort, FY2009-10 to FY2019-20

Primary outcome: accepted workers' compensation claims for traumatic finger or thumb amputation (TOOCS Nature of Injury 139, Bodily Location 461/462) per 100,000 covered jobs. (A) 2012 adoption cohort (NSW, QLD, ACT, NT) and comparators. (B) 2013 cohort (SA, TAS) and comparators. Solid lines show jurisdictions in the treated cohorts; dashed and dotted lines show the comparators (VIC, never treated; WA, not yet treated within this window). The vertical grey line marks each cohort's own mixed-exposure financial year. Gaps indicate financial years suppressed by Safe Work Australia (fewer than five claims); these are not interpolated. No single jurisdiction shows an isolated, cohort-specific pre-treatment trend break coinciding with its own implementation date

NSW: New South Wales; QLD: Queensland; ACT: Australian Capital Territory; NT: Northern Territory; SA: South Australia; TAS: Tasmania; VIC: Victoria; WA: Western Australia; WHS: Work Health and Safety; TOOCS: Type of Occurrence Classification System

Primary outcome

The pooled post-treatment incidence-rate ratio was 1.03 (descriptive 95% CI: 0.85-1.25; naïve cluster-robust, see Materials and Methods) across all jurisdictions, and 1.00 (0.83-1.20) excluding NSW. Under exact randomization inference, neither estimate provided evidence of an association (all jurisdictions: RI-β p = 0.89, RI-t p = 0.88; excluding NSW: RI-β p = 0.98, RI-t p = 0.98). A naïve cluster-robust pretrend test was statistically significant (χ^2=10.94, p=0.012, all jurisdictions; χ^2 = 47.42, p < 0.001 excluding NSW), which would ordinarily raise concerns about differential pretreatment trends, but the corresponding exact randomization test did not detect differential pretreatment trends (p = 0.48). With covariate adjustment, the pooled IRR was 1.08 (0.89-1.31), and 1.02 (0.85-1.22) excluding NSW; the adjusted naïve pretrend test was not significant (p = 0.44). This does not establish that parallel trends held, particularly given the short pre-period discussed in Limitations (Table 3).

Outcome/specification	Jurisdictions (n)	IRR (95% CI)	RI-β p-value	RI-t p-value	Pretrend p (naïve)	Pretrend RI p-value
Primary digit amputation – all jurisdictions	8	1.03 (0.85–1.25)	0.89	0.88	0.012	0.48
Primary digit amputation – excl. NSW	7	1.00 (0.83–1.20)	0.98	0.98	< 0.001	–
Machinery-related digit amputation – all	7	0.99 (0.80–1.23)	0.96	0.97	< 0.001	0.27
Machinery-related digit amputation – excl. NSW	6	0.96 (0.77–1.20)	–	–	< 0.001	–
Broader machinery-related time-loss – all	8	0.99 (0.92–1.06)	0.80	0.81	< 0.001	0.91
Broader machinery-related time-loss – excl. NSW	7	0.97 (0.88–1.07)	–	–	< 0.001	–

TABLE 3: Pooled post-treatment estimates and randomization inference by outcome

CIs are naïve cluster-robust intervals, reported descriptively because of the small number of clusters; inferential conclusions are based on RI-β and RI-t. All values within each row derive from the same bare model (jurisdiction and financial-year fixed effects with a log(jobs) offset). Jurisdictions (n) is the number of jurisdictions contributing data; machinery-related outcomes exclude ACT because all its cells were suppressed. Covariate-adjusted results are in Table 4 and the text (machinery-related and broader time-loss outcomes: Supplementary File, Table S3; see Google Doc link in the Appendices)

–: not computed. IRR: incidence-rate ratio; CI: confidence interval; RI-β: coefficient-based randomization inference (420 relabelings); RI-t: t-statistic-based randomization inference; Pretrend p (naïve): cluster-robust Wald test of the pre-adoption event-study coefficients; Pretrend RI p-value: its randomization-inference counterpart

Secondary outcomes

For machinery-related digit amputation, the bare-model pooled IRR was 0.99 (descriptive 95% CI: 0.80-1.23) across all jurisdictions and 0.96 (0.77-1.20) excluding NSW; RI-β p = 0.96, RI-t p = 0.97. The naïve pretrend test was significant ($\chi^2 = 127.56$, $p < 0.001$), but the randomization test was not (p = 0.27). For the broader machinery-related time-loss hand-injury outcome, the bare-model pooled IRR was 0.99 (0.92-1.06) overall and 0.97 (0.88-1.07) excluding NSW; RI-β p = 0.80, RI-t p = 0.81. Again, the naïve pretrend test was significant ($\chi^2 = 36.01$, $p < 0.001$), whereas the randomization test was not (p = 0.91). Machinery-related outcomes were estimable in seven jurisdictions because all ACT machinery-related cells were suppressed (Table *3*). With covariate adjustment, the machinery-related IRR was 1.05 (0.81-1.36) overall and 0.99 (0.69-1.41) excluding NSW, and the broader time-loss IRR was 1.03 (0.93-1.14) overall and 1.02 (0.87-1.18) excluding NSW (Supplementary File, Table S3).

Event-study estimates

In event-study estimates (reference period: two years before adoption), pre-adoption coefficients were 0.85 (descriptive 95% CI: 0.59-1.22) three years before adoption in the 2012 cohort, and 0.71 (0.51-0.99) and 0.90 (0.66-1.23) four and three years before adoption in the 2013 cohort. Post-adoption coefficients ranged from 0.70 to 1.32 without a sustained monotonic decline in either cohort. The naïve intervals for the 2013 cohort's four-year lead (0.71) and its four-year post-adoption coefficient (0.70, 0.58-0.83) both excluded 1.0; with only two jurisdictions in that cohort, these intervals are unreliable, and neither pattern was sustained in the adjacent years (Supplementary File, Table S2). The joint pretrend tests are reported above (Table *3*).

Triple-difference: machinery versus non-machinery claims

An initial triple-difference specification without mechanism-by-year fixed effects suggested a larger post-treatment decline in machinery-related than non-machinery claims (IRR: 0.84, descriptive 95% CI: 0.71-0.99); however, all 420 treatment-timing permutations produced a negative RI-β coefficient and the RI-t permutation distribution was itself strongly negative (mean t = -4.57), a pattern more consistent with an unmodeled nationwide, calendar-time-varying difference between machinery and non-machinery claim rates than with an association specific to WHS treatment timing. I therefore added mechanism-by-year fixed effects to absorb any such secular mechanism-specific trend. Under this fuller specification, the triple-difference estimate was null (IRR: 0.99, descriptive 95% CI: 0.83-1.17) and was not supported under exact randomization inference (RI-β p = 0.94; RI-t t = -0.17, p = 0.93), consistent with the primary and machinery-specific results above. These analyses rest on the seven jurisdictions with machinery-related data.

Sensitivity analyses

Sensitivity analyses are summarized in Table 4, in the bare model and, where indicated, with covariate adjustment. The primary inference under randomization inference was unchanged. Excluding NSW gave IRRs of 1.00 (bare) and 1.02 (adjusted). Restricting to the six jurisdictions with negligible suppression gave 1.01 and 1.08. Truncating follow-up at FY2015-16 gave 0.98 (0.76-1.25) and 0.97 (0.83-1.14), and extending it through FY2018-19 gave 1.00 (0.83-1.22) and 1.04 (0.86-1.26). Replacing suppressed cells with 0 or 4 gave 1.01 and 1.02, and adjustment for industry composition alone gave 1.08 (0.92-1.27). The primary and six-jurisdiction covariate-adjusted estimates both round to 1.08 (0.89-1.31) but differ at the third decimal (1.084 and 1.077; Table 4). All confidence intervals in this paragraph are naïve cluster-robust and descriptive only.

Specification	Bare model IRR (95% CI)	Covariate-adjusted IRR (95% CI)	Randomization inference p
Primary analysis, eight jurisdictions	1.029 (0.845–1.252)	1.084 (0.894–1.313)	RI-β 0.89; RI-t 0.88
Excluding NSW	0.996 (0.826–1.200)	1.019 (0.855–1.216)	RI-β 0.98; RI-t 0.98
Six jurisdictions (excluding NT and ACT)	1.014 (0.834–1.233)	1.077 (0.889–1.306)	–
Industry-composition covariates only	–	1.079 (0.919–1.265)	–
Follow-up truncated at FY2015-16	0.977 (0.762–1.251)	0.972 (0.827–1.143)	–
Follow-up through FY2018-19	1.005 (0.828–1.220)	1.044 (0.863–1.262)	–
Suppressed cells replaced with 0	1.013 (0.844–1.218)	–	–
Suppressed cells replaced with 4	1.021 (0.848–1.229)	–	–
Excluding Victoria (leave-one-out)	1.158 (1.013–1.325)	1.160 (1.049–1.283)	RI-β 0.62
Clean-control stacked (cohort-specific)†	1.041 (0.875–1.238)	–	–
Triple-difference, machinery vs non-machinery, with mechanism-by-year fixed effects*	0.985 (0.830–1.170)	0.982 (0.825–1.170)	RI-β 0.94; RI-t 0.93
Western Australia vs Victoria only, primary outcome, FY2016-17 to FY2023-24‡	0.640 (0.569–0.720)	–	–

TABLE 4: Sensitivity and exploratory analyses in the bare model and with covariate adjustment

*Both triple-difference models include mechanism-by-year fixed effects and rest on seven jurisdictions with machinery-related data; a model without them gave a bare IRR of 0.838 (0.707–0.993), with all 420 relabelings negative (RI-t mean t = −4.57). †Each treated cohort compared only with VIC and WA using stratum-specific jurisdiction and financial-year fixed effects; one sparsely populated stratum-year cell produced a rank-deficiency warning (Supplementary File, Section A; see Google Doc link in the Appendices). ‡Illustrative only: two jurisdictions, cluster-robust inference degenerate

The primary analysis and the six-jurisdiction model with covariates both round to 1.08 (0.89–1.31) at two decimals but differ at three (1.084 [0.894–1.313] and 1.077 [0.889–1.306]); the six-jurisdiction model omits NT and ACT (62 versus 73 observations)

IRR: incidence-rate ratio; CI: confidence interval; IRRs and 95% CIs are shown to three decimals. Bare model: jurisdiction and financial-year fixed effects with a log(jobs) offset. Covariate-adjusted model: additionally adjusted for the manufacturing, construction, and agriculture shares of employment, the unemployment rate, and the female share of employment. Intervals are naïve cluster-robust (descriptive); heteroskedasticity-consistent (HC1) standard errors were used for the two-jurisdiction comparison. Randomization inference (RI) uses the bare model (RI-β, coefficient-based; RI-t, t-statistic-based); –: not computed or not applicable

Leave-one-jurisdiction-out re-estimation in the bare model gave IRRs from 0.96 (excluding WA) to 1.16 (excluding VIC). Excluding Victoria produced the only naïve cluster-robust confidence interval that excluded 1.0 (1.16, descriptive 95% CI: 1.01-1.32); this was not supported by exact randomization inference on the

remaining seven jurisdictions (RI-β p = 0.62). In the covariate-adjusted model, the exclusions gave IRRs from 1.00 to 1.20, and excluding Victoria again gave 1.16 (1.05-1.28). These intervals are descriptive; inference rests on randomization inference.

To directly test whether comparisons involving already-treated jurisdictions as controls materially drove the pooled result, I fit a cohort-specific stacked specification comparing the 2012 and 2013 cohorts only against Victoria and Western Australia. The resulting estimate (IRR: 1.04, descriptive 95% CI: 0.88-1.24) changed by approximately one percentage point from the pooled specification, a pattern that also held with suppressed cells replaced by 0 (stacked IRR: 1.02) or 4 (1.03), suggesting that already-treated-as-control comparisons did not materially drive the primary result (a rank-deficiency warning from one sparsely populated stratum-year cell is explained in Supplementary File, Section A; see Google Doc link in the Appendices).

Restricting to WA and Victoria over FY2016-17 to FY2023-24 (provisional), the pooled primary-outcome IRR was 0.64 (descriptive 95% CI: 0.57-0.72) and the machinery-related IRR was 0.51 (0.44-0.59), whereas the broader time-loss IRR was 1.03 (0.92-1.16). With only two jurisdictions, cluster-robust inference is degenerate, and these estimates are illustrative only; the most recent financial year (FY2023-24) is also provisional, so incomplete claims maturation may contribute to the apparent reduction. A naïve Wald pretrend test returned p = 0.92, but the covariance matrix was rank-deficient given only two jurisdictions (Supplementary File, Section A; see Google Doc link in the Appendices); no exact-randomization version exists for a two-jurisdiction comparison, so this diagnostic should be treated as unreliable, not reassuring.

Statistical power

In the post hoc simulation, exact RI-β had 6% power with Poisson counts and 2% with overdispersed counts to detect a true 10% reduction (IRR: 0.90), 21% and 9% for a 20% reduction, 48% and 28% for a 30% reduction, and 79% and 50% for a 40% reduction. Power of 80% required a true IRR of about 0.59 with Poisson counts and about 0.44 with overdispersed counts, corresponding to reductions of roughly 40% to 55% (Supplementary File, Section B7, Table S5; see Google Doc link in the Appendices). With no true effect, the test was rejected in fewer than 1% of simulations, indicating that it is conservative.

Discussion

I found no evidence that harmonizing Work Health and Safety laws was associated with reduced traumatic occupational digit amputation among the primary 2012 and 2013 adoption cohorts. The primary IRR was 1.03 (descriptive 95% CI: 0.85-1.25) across all jurisdictions and 1.00 (0.83-1.20) excluding NSW, with no evidence of an association under exact randomization inference (RI-β p = 0.89 overall and p = 0.98 excluding NSW). The primary inference under randomization inference was unchanged across sensitivity analyses.

NSW deserves separate attention because of a specific competing policy change. About six months after NSW's WHS harmonization date, on June 19, 2012, a separate NSW workers' compensation benefit-scheme reform took effect, tightening eligibility and introducing time limits on income support. Using the same Safe Work Australia data underlying the outcome used here, an interrupted time series evaluation, which did not address WHS harmonization, found this reform reduced accepted NSW claims by 15.3% relative to seven other jurisdictions [10]. This is a relevant competing explanation for any claims-based change in NSW during this period. Excluding NSW left the result materially unchanged (IRR: 1.00 versus 1.03), and exact randomization inference again showed no evidence of an association (RI-β p = 0.98, RI-t p = 0.98).

A secondary contribution is methodological. With only seven or eight jurisdictions, conventional cluster-robust standard errors performed poorly: a naïve pretrend test was significant (p = 0.012, all jurisdictions; p < 0.001 excluding NSW), which would ordinarily indicate a parallel-trends violation, but was not supported under exact randomization inference (p = 0.48). The same pattern recurred for both machinery-specific outcomes, consistent with literature showing conventional cluster-robust inference can over-reject with few clusters [12,13] and that randomization-inference procedures can be useful when their exchangeability assumptions are credible [14,15]. Because coefficient-based randomization inference may itself be poorly calibrated with heterogeneous cluster sizes [14], I also implemented a t-statistic-based alternative (RI-t), which closely agreed with RI-β for all three principal outcomes.

The main message is straightforward. WHS harmonization did not introduce workplace safety law into previously unregulated jurisdictions - every jurisdiction already had some occupational health and safety (OHS) statute well before 2012. Harmonization instead standardized core statutory duties, definitions, and penalty frameworks across jurisdictions that were already regulated.

Digit amputation was also already declining. Victoria never adopted the model WHS laws, yet its own age-standardized hospital admission rate for work-related digit amputation fell from 12.54 to 4.71 per 100,000 employed persons between 2005/06 and 2022/23 [3]; the Victorian workers' compensation claim rate in this study also fell, from 7.0 to 5.6 per 100,000 jobs between FY2009-10 and FY2019-20 (Table *1*). Plausible drivers include a shrinking, more mechanized manufacturing sector and increasing automation, patterns identified in the United States (US) [16] and elsewhere [17-20]. A comparable pattern has been shown for a federal safety standard: a rigorous evaluation of the US lockout/tagout rule found no detectable effect on

machinery-related fatal injury, attributed to voluntary adoption before the standard took effect combined with persistently low compliance [21], and survey evidence shows guarding and lockout/tagout compliance gaps persist regardless of statutory framework [22].

This offers one explanation for my findings: harmonization may have reached a hazard that was already regulated and already improving, leaving limited room for a further legal change to produce another measurable decline. This does not mean occupational safety regulation does not work. It means I did not detect an additional reduction from this specific reform against an already-declining background.

The study had low power to detect modest effects. In a post hoc simulation based on the observed panel, exact randomization inference had 80% power only for a true reduction of roughly 40% to 55% (IRR about 0.59 with Poisson counts and 0.44 allowing for the overdispersion in the data) and less than 50% power for a 30% reduction (see Results). The null result therefore excludes only very large reductions and does not exclude the modest reductions that a reform of this kind might plausibly produce. The descriptive 95% interval (0.85-1.25) is much narrower than this and should not be read as excluding reductions greater than 15%, because it is unreliable with so few clusters. The test was also conservative in simulation, rejecting rarely when there was no true effect, so the reported p-values should be read as conservative. The hazard-maturity explanation above is one hypothesis that these data cannot test directly; others include low statistical power, offsetting changes in claim reporting or acceptance, effects too small or too slow to detect in aggregate counts, and concurrent reforms.

Machinery-related outcomes were likewise null (IRR: 0.99, $p = 0.96$; broader time-loss IRR: 0.99, $p = 0.80$). An initial triple-difference specification suggested a larger decline in machinery-related claims, but the same pattern appeared across all 420 treatment-timing relabelings, a signature of a nationwide machinery-specific trend rather than an effect of WHS timing. With mechanism-by-year fixed effects, the estimate was null (IRR: 0.99; RI-β $p = 0.94$; RI-t $p = 0.93$).

Western Australia's 2022 adoption showed lower primary-outcome (IRR: 0.64) and machinery-related (IRR: 0.51, descriptive 95% CI: 0.44-0.59) rates relative to Victoria, but no difference in the broader time-loss outcome (IRR: 1.03, 0.92-1.16). I do not consider these findings confirmatory: they rest on a two-jurisdiction comparison with degenerate cluster-robust inference, an unreliable pretrend diagnostic (see Results), and a most-recent year that was still provisional. I regard this analysis as exploratory.

These findings should not be interpreted as evidence that occupational safety regulation is ineffective. Rather, given the limited power described above, they concern the marginal effect of statutory harmonization in a setting where the relevant hazard was already regulated and declining. A hypothesis for future study is that comparable regulatory reforms may produce larger measurable effects in emerging hazards that have not undergone similar secular declines; the present data cannot test that possibility.

The outcome relies on accepted workers' compensation claims, not on every injury that occurred. Reporting, claim submission, and claim acceptance can all change independently of true injury rates. A Michigan study found employers reported only about 44% of work-related hospitalizations covered by a mandatory reporting requirement, with the worst compliance in small employers and in industries most relevant to machinery injury [23]. I chose digit amputation because it is severe and likely to have relatively high ascertainment [24]. Even so, I cannot exclude that WHS harmonization's changes to definitions and penalties altered claims or reporting behavior independently of any true change in injury risk.

Strengths and limitations

This study has several strengths: near-national administrative data across all eight jurisdictions, two adoption cohorts within the same analysis window, and explicit treatment of the statistical problems created by having very few clusters. The primary and clean-control stacked estimates also agree closely (1.03 versus 1.04), making it less likely that comparisons involving already-treated jurisdictions explain the main result.

However, this study has several limitations as well. First, I used aggregate jurisdiction-level claims rather than individual-level data. Second, Safe Work Australia suppressed cells with fewer than five claims, concentrated in NT and ACT. Alternative assumptions about these cells produced similar estimates. Third, the 2012 and 2013 cohorts had only two to three fully pre-adoption years. This sharply limits any pretrend diagnostic's ability to assess the parallel-trends assumption. A clean result indicates no differential pretrend was detected, not that parallel trends are established. Fourth, I could not estimate the permanently disabling machinery-related outcome at the jurisdiction level in the primary analysis; jurisdiction-level data for this outcome has since become available, and analysis is planned as future work. I also could not isolate which specific element of harmonization might drive an undetected effect.

Fifth, the primary model used a static two-way fixed-effects design, which can in principle place implicit weight on already-treated cohorts as controls for later-treated cohorts [11]. A clean-control stacked sensitivity produced a materially similar estimate, though this does not remove all concerns about staggered treatment timing. Sixth, randomization inference assumes that treatment-timing labels are exchangeable

across jurisdictions, which cannot be verified, and coefficient-based randomization inference can be poorly calibrated with heterogeneous cluster sizes. A closely agreeing t-statistic-based check using the same sample addressed the second concern but not the first; the p-values should be read as conditional on the exchangeability assumption. Seventh, I rely on naïve confidence intervals as descriptive summaries rather than an interval inverted from the randomization test.

Eighth, exposure was coded as a binary indicator of statutory adoption, which does not capture implementation intensity, enforcement, compliance, or delayed effects, and these may have differed across jurisdictions. A published evaluation of the harmonization reported that the number of inspectors per 10,000 employees stayed roughly constant in all but the smallest jurisdictions (ACT and NT), that proactive inspections rose in NSW and fell in QLD after the reform, and that total court-awarded fines fell in most jurisdictions between 2011-12 and 2013-14 [9]. Differences of this kind could dilute or mask an association with adoption. The event-study estimates allow for delayed effects but cannot separate them from the absence of an effect.

Ninth, other jurisdiction-specific changes may have coincided with adoption. Beyond the 2012 NSW benefit-scheme reform discussed above, Safe Work Australia's scheme histories record, for example, a 2013 permanent-impairment threshold for damages claims in Queensland and the replacement of South Australia's scheme by the Return to Work scheme on July 1, 2015 [25]. I did not compile a complete timeline of enforcement or compensation-scheme changes for every jurisdiction, and such changes could bias the estimates in either direction. Similarly, compensation coverage, eligibility, and acceptance practices may have changed within jurisdictions over time. Leave-one-out and NSW-exclusion analyses did not change the inference under randomization inference (Table *4*). Tenth, with only eight jurisdictions, the study had low statistical power: in simulation, exact randomization inference reliably detected only large reductions (see Results), so the null finding does not exclude modest effects. Finally, Australia's specific regulatory architecture may limit generalizability to other countries.

Conclusions

In this natural experiment, I found no evidence that harmonizing Work Health and Safety laws reduced accepted workers' compensation claims for traumatic digit amputation among the primary adoption cohorts, either overall or for machinery-related mechanisms. The study had low statistical power, however, and could reliably detect only large reductions (roughly 40% to 55% or more); therefore, modest effects cannot be ruled out. This finding should not be interpreted as evidence that workplace safety laws do not matter. One possible interpretation is that, by the time harmonization was implemented in each jurisdiction, digit amputation was already a declining and at least partially regulated hazard, leaving limited scope for a statutory change of this kind to further reduce incidence. However, low statistical power and the other explanations discussed above remain plausible alternatives. Whether regulatory reform has greater potential to influence less mature hazard categories is a question that cannot be answered using these data.

Appendices

Supplementary file link

https://docs.google.com/document/d/1FFC-XNy0YK6EJxLCZgVV7MWFAUwZhMxWPKKzXGTfEK4/edit?usp=sharing

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Michael J. Franco

Acquisition, analysis, or interpretation of data: Michael J. Franco

Drafting of the manuscript: Michael J. Franco

Critical review of the manuscript for important intellectual content: Michael J. Franco

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might

have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

Data availability. Data are available on reasonable request. The Safe Work Australia custom data extract (Submission 875), the constructed analytic panel, and the analysis code are available from the author (Michael J. Franco, franco-michael@cooperhealth.edu) on reasonable request. The Australian Bureau of Statistics Labour Force data used to construct covariates are publicly available from the ABS. I thank Safe Work Australia for providing the custom data extract (Submission 875). I am the sole author. I conceived the research question, designed the study and the outcome and cohort definitions, obtained the data from Safe Work Australia, directed the analyses, interpreted the results, and wrote, reviewed, and edited the manuscript. Claude (Anthropic) was used to draft and cross-check portions of the Python analysis code, verify citations, and prepare the tables and figures. ChatGPT (OpenAI) was employed to independently audit the analysis pipeline and review the draft, and some Discussion wording was adapted from its suggestions. The regulatory-timing interpretation in the Discussion and Conclusions is mine. No AI system is an author, and I take full responsibility for the content.

References

1. Crowe CS, Massenburg BB, Morrison SD, et al.: Global trends of hand and wrist trauma: a systematic analysis of fracture and digit amputation using the Global Burden of Disease 2017 Study. Inj Prev. 2020, 26:i115-24. 10.1136/injuryprev-2019-043495
2. Cuan X, Zhang J, Lu Z, et al.: Global epidemiology of amputation-related disability from 1990 to 2021: geographic variation and projections to 2050. J Cachexia Sarcopenia Muscle. 2026, 17:e70356. 10.1002/jcsm.70356
3. Wah W, Berecki-Gisolf J, Walker-Bone K: Work-related traumatic finger and thumb amputations: 20-year trends. Occup Med (Lond). 2026, 76:57-63. 10.1093/occmed/kqaf134
4. Panagopoulou P, Antonopoulos CN, Dessypris N, Kanavidis P, Michelakos T, Petridou ET: Epidemiological patterns and preventability of traumatic hand amputations among adults in Greece. Injury. 2013, 44:475-80. 10.1016/j.injury.2012.10.008
5. Mayrhofer-Schmid M, Aman M, Stolle A, et al.: Analysis of epidemiology, etiology and injury patterns in 2,179 digit amputations. Sci Rep. 2025, 15:32657. 10.1038/s41598-025-18983-y
6. Safe Work Australia: Evaluation Plan for the Harmonisation of Work Health and Safety in Australia. Safe Work Australia, Canberra, Australia; 2011. https://www.safeworkaustralia.gov.au/system/files/documents/1702/final-evaluation-plan-whs-harmonisation.pdf.
7. Gunningham N: Impacts of work health and safety harmonisation on very large businesses . SSRN Electron J. 2016, 4:23-5. 10.2139/ssrn.2848566
8. Barrett R, Mayson S, Bahn S: Small firms and health and safety harmonisation: potential regulatory effects of a dominant narrative. J Ind Relat. 2014, 56:62-80. 10.1177/0022185613498629
9. Bilgrami A, Cutler H, Sinha K: Higher penalties, broader definitions, and national standards: did harmonized Australian workplace health and safety laws reduce workers' compensation receipt?. Ind Relat. 2025, 64:77-102. 10.1111/irel.12357
10. Collie A, Beck D, Gray SE, Lane TJ: Impact of legislative reform on benefit access and disability duration in workers' compensation: an interrupted time series study. Occup Environ Med. 2020, 77:32-9. 10.1136/oemed-2019-106063
11. Goodman-Bacon A: Difference-in-differences with variation in treatment timing . J Econom. 2021, 225:254-77. 10.1016/j.jeconom.2021.03.014
12. Cameron AC, Gelbach JB, Miller DL: Bootstrap-based improvements for inference with clustered errors . Rev Econ Stat. 2008, 90:414-27. 10.1162/rest.90.3.414
13. Canay IA, Santos A, Shaikh AM: The wild bootstrap with a "small" number of "large" clusters . Rev Econ Stat. 2021, 103:346-63. 10.1162/rest_a_00887
14. MacKinnon JG, Webb MD: Randomization inference for difference-in-differences with few treated clusters . J Econom. 2020, 218:435-50. 10.1016/j.jeconom.2020.04.024
15. Conley TG, Taber CR: Inference with "difference in differences" with a small number of policy changes . Rev Econ Stat. 2011, 93:113-25. 10.1162/REST_a_00049
16. Reavey PL, Stranix JT, Muresan H, Soares M, Thanik V: Disappearing digits: analysis of national trends in amputation and replantation in the United States. Plast Reconstr Surg. 2018, 141:857e-67e. 10.1097/PRS.0000000000004368
17. Saari J: Long-term development of occupational accidents in Finland . Scand J Work Environ Health. 1982, 8:85-93. 10.5271/sjweh.2484
18. Subramanian A, Desai A, Prakash L, Mital A, Mital A: Changing trends in US injury profiles: revisiting non-fatal occupational injury statistics. J Occup Rehabil. 2006, 16:123-55. 10.1007/s10926-005-9012-1
19. Arana-Landín G, Laskurain-Iturbe I, Iturrate M, Landeta-Manzano B: Assessing the influence of Industry 4.0 technologies on occupational health and safety. Heliyon. 2023, 9:e13720. 10.1016/j.heliyon.2023.e13720
20. Bylsma VF, Webber BJ, Erich RA, Voss JD: Musculoskeletal injuries and automation in aerial port operations . Aerosp Med Hum Perform. 2020, 91:669-73. 10.3357/AMHP.5475.2020
21. Bulzacchelli MT, Vernick JS, Webster DW, Lees PS: Effects of the Occupational Safety and Health Administration's control of hazardous energy (lockout/tagout) standard on rates of machinery-related fatal occupational injury. Inj Prev. 2007, 13:334-8. 10.1136/ip.2007.015677
22. Parker DL, Yamin SC, Brosseau LM, Xi M, Gordon R, Most IG, Stanley R: National Machine Guarding Program: Part 1. Machine safeguarding practices in small metal fabrication businesses. Am J Ind Med. 2015,

58:1174-83. 10.1002/ajim.22514

23. Reilly MJ, Wang L, Rosenman KD: Evaluation of the characteristics of workers injured on the job requiring hospitalization, and employer compliance with OSHA's reporting requirement for these work-related hospitalizations. Am J Ind Med. 2023, 66:109-21. 10.1002/ajim.23447
24. Sears JM, Bowman SM, Hogg-Johnson S: Using injury severity to improve occupational injury trend estimates. Am J Ind Med. 2014, 57:928-39. 10.1002/ajim.22329
25. Safe Work Australia: Comparison of Workers' Compensation Arrangements in Australia and New Zealand 2021 (28th Edition). Safe Work Australia, Canberra, Australia; 2021. https://www.safeworkaustralia.gov.au/book/comparison-workers-compensation-arrangements-australia-and-new-zealand-2021....