Causal Investigation and Analysis of Crane Incident Reports

A hybrid of language-model extraction, causal discovery, and physics simulation

ANNUAL MEETING OF THE CRANE SAFETY RESEARCH CENTER

Presenter: Yutian Pang, Ph.D.

Postdoctoral Researcher, The University of Texas at Austin

Oct 23rd, 2026

Three Questions about Fatal Crane Accidents

RQ1	Can a language model turn free-text reports into reliable causal chains?	Hard because a cause often sits in one clause of a long narrative.
RQ2	Which factors cause each injury outcome, and which decide whether it is fatal?	Hard because confounders and chained causes distort raw counts.
RQ3	How likely is a given lift to fail, and which safety measure helps most?	Hard because the reports contain no lift that went well.

Data: 7,188 OSHA and Vertikal reports, of which 5,723 describe crane accidents and 49% of those are fatal.

> Each question needs its own tool: a language model, a causal graph, and a physics simulator.

0. WORKFLOW

A Hybrid Approach in Three Layers

LAYER 1

Language Model Extraction

A local language model reads each report and fills the crane_v1 schema with the causal chain.

Output
7,177 chains over 55 factor types

LAYER 2

Causal Discovery and Estimation

PC, GES, and FCI learn one graph per injury outcome. Backdoor adjustment gives interventional effects.

Output
P(Death | Fail)

LAYER 3

Physics Simulation

A calibrated tipover simulator gives the chance that a lift fails, which the reports cannot give.

Output
P(Fail | Setup)

> The data layers say what a failure leads to. The physics layer says how likely the failure is.

0. DATA VIEW

What do the two report sources see?

OSHA Records the Victim, Vertikal Records the Machine

Source	Reports	Years	Median words	Content
OSHA IMIS summaries	4,340	1973 to 2025	84	Worker injuries and deaths in a fixed template, with a fatality flag and victim records
Vertikal.net news	2,848	2001 to 2026	163	Machine-focused stories, damage-only events, cause updates, and industry news

GROUND TRUTH FOR SEVERITY

OSHA codes the degree of injury for 5,243 victims, independent of any model.

CHECK ON EXTRACTION

Human labels on 1,219 OSHA reports test the language model later.

NOISE TO REMOVE

About a quarter of Vertikal stories are opinion, updates, or industry news.

> The sources report different kinds of accident, so the source enters every causal adjustment.

Data: OSHA, Accident Investigation Search (IMIS) · Vertikal.net news archive

0. DATA VIEW

One Dataset Record, OSHA Record 3521

At approximately 10:35 a.m. on October 19, 1990, Employee #1, a carpenter, was directing a 15-ton P and H Omega 15 hydraulic mobile crane equipped with a 62 ft long telescopic boom as it hoisted a **1,900 lb concrete form panel**. The 7 ft by 16 ft form was constructed of plywood with steel backing. According to the lift charts, when the crane was upright and the boom angle at 40 degrees, with 60 ft of boom extended and an approximate boom radius of 45 ft, **the crane could lift a load of 500 lb**. **With the outriggers extended, the lift chart stated that 5,100 lb could be handled**. During hoisting, **the crane tipped over on its left side**. **The box tube boom struck Employee #1**, killing him.

Structured Fields

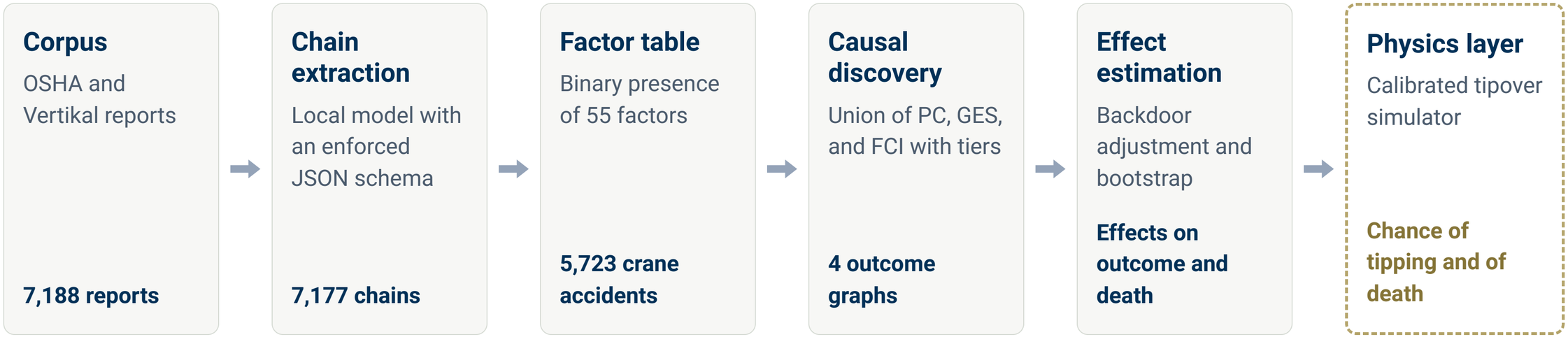
Field	Value
Source	OSHA IMIS
Year	1990
Length	128 words
Fatality flag	Yes
Degree of injury	Fatality
Inspection number	111914495

Highlighted phrases become the chain nodes, evidence quotes, and drivers in the schema output. The fatality flag is never shown to the model as a label.

Source: OSHA, Accident Investigation Search (IMIS), inspection 111914495

0. WORKFLOW

The Pipeline from Reports to Risk



MODEL

Qwen3.6-35B-A3B served by vLLM, run of 23 September 2026

SCHEMA

crane_v1, three levels, closed vocabularies

CASE STUDY

Tipover, with a structural collapse simulator in development

> Every step keeps its counts, so each number in the talk traces back to the reports.

1. INFORMATION EXTRACTION

The Extraction Schema Has Three Levels

Record Level

One per report

- Record type
- Equipment class (BLS OIICS)
- Crane state
- Primary mechanism
- Severity, six-point scale
- Drivers: load, chart rating, radius, boom, wind, outrigger state

Node Level

One per factor in the chain

- Factor type, one of 55
- Class: condition, event, or outcome
- Operation phase
- Cause role: primary, contributing, outcome
- Trigger phrase

Link Level

One per cause and effect pair

- Cause node, always earlier in the list, so no loops
- Strength: direct or contributing
- Evidence quote from the report

> The server enforces every closed vocabulary during decoding with a JSON Schema. A validator checks the remaining rules, with one repair request per failed output.

BLS (2012), OIICS Manual 2.01 · Raviv, Fishbain, and Shapira (2017), Safety Science 91

55 Factor Types from Organization to Outcome

Layer	Types	Examples	Grounded in
Organization	5	Lift planning, supervision, inspection, safety management	HFACS, Shapira et al.
Preconditions	3	Untrained operator, fatigue, blind lift	Raviv et al. (2017)
Environment	8	Wind, soft ground, power line proximity	Subpart CC, FEM 5.016
Equipment	7	Wire rope, rigging hardware, structural component failure	Raviv et al., Subpart CC
Unsafe acts	17	Capacity exceeded, outriggers not deployed, worker in hazard zone	Subpart CC, CPWR
Mechanism	9	Tipover, structural collapse, load drop, power line contact	WORM bow tie center
Outcome	6	Electrical exposure, struck by object, caught in, fall	BLS OIICS event codes

> Each type becomes one binary variable. The layer order follows HFACS and the WORM bow tie, and it becomes the tier order for causal discovery.

Shappell and Wiegmann (2000), HFACS · Aneziris et al. (2008), Safety Science 46 · Raviv et al. (2017), Safety Science 91 · BLS (2012), OIICS · OSHA 29 CFR 1926 Subpart CC · FEM 5.016 (2017) · McCann et al. (2009), CPWR

Schema Output for Record 3521

RECORD LEVEL

Mobile telescopic crane, in service and lifting. Primary mechanism tipover. Severity fatal. Drivers: load 1,900 lb, rating 500 lb, radius 45 ft, boom 60 ft, angle 40°.

NODE LEVEL

#	Factor type	Role	Caused by
0	Capacity exceeded	Primary	none
1	Outriggers not properly deployed	Primary	none
2	Loss of stability, tipover	Contributing	0, 1
3	Struck by object	Outcome	2

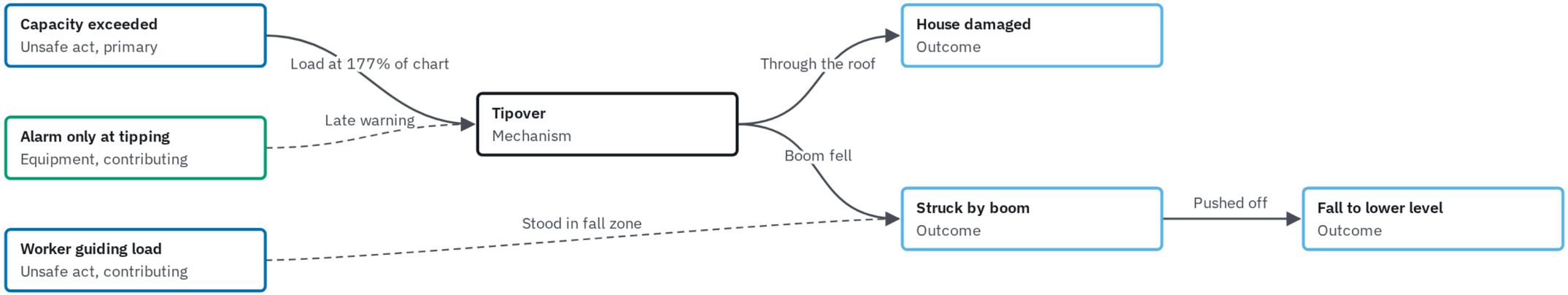
NODE 2 OF THE CHAIN, AS RETURNED


```
{
  "idx": 2,
  "class": "event",
  "factor_type": "LOSS_OF_STABILITY_TIP_OVER",
  "trigger": "crane tipped over on its left side",
  "phase": "hoisting",
  "cause_role": "contributing",
  "caused_by": [
    {"src": 0, "strength": "direct", "evidence":
"1,900 lb load lifted where the crane could lift a
load of 500 lb"},
    {"src": 1, "strength": "direct", "evidence": "With
the outriggers extended, the lift chart stated that
5,100 lb could be handled"}]
}
```

1. INFORMATION EXTRACTION

What does one extracted chain look like?

The Chain Separates Act, Mechanism, and Outcomes



OSHA record 3237, a truck crane tipover. Solid arrows are direct links and dashed arrows are contributing links. Extracted drivers: load 4,303 lb, chart rating 2,425 lb (177% of the rating), radius 58 ft, boom 92 ft.

> Each node becomes one causal variable, and the extracted drivers feed the physics layer.

Can we trust the extracted chains?

Agreement Is High for Wind, and Most Disagreements Are Label Errors

The Run

- Valid chain for 7,177 of 7,188 reports
- 94 outputs (1.3%) needed the one repair request
- 415 records removed as not accident reports, mostly industry news
- About 4.4 reports per second on four GPUs

Agreement with Human Labels

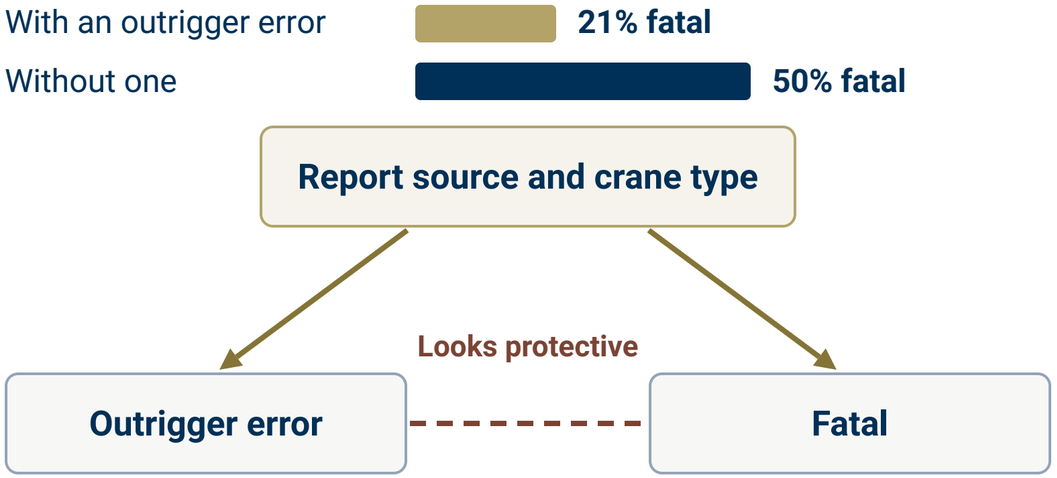
Cohort	Precision	Recall	Kappa
Wind	0.77	0.81	0.78
Mobile crane	0.85	0.68	0.65
Static crane	0.70	0.57	0.52

On review, most crane-type disagreements are human label errors, such as a crawler crane labeled static.

> Test runs exposed three schema problems, all fixed before the final run: crane involvement, ambiguous wind values, and record type for OSHA.

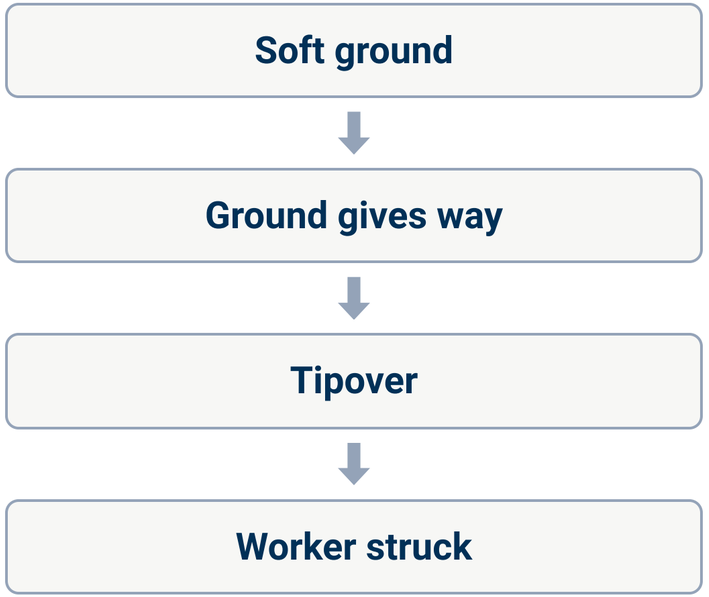
Why Simple Counts Mislead

Do Outrigger Errors Save Lives?



59% of these accidents come from Vertikal, which mostly reports damage-only tipovers. Adjusted for source, crane type, and the 2015 rule, the effect is −0.09 (95% interval −0.22 to 0.01), not significant.

Causes Act in Sequence



Counted one by one, soft ground, ground failure, and tipover each get credit for the same injury. The chain shows that soft ground acts through the later steps.

> A causal graph handles both. It adjusts for common causes such as source and crane type, and it keeps the steps of a chain in order.

Pearl (2009), Causality, 2nd ed., Cambridge University Press

One Graph per Injury Outcome

Causes → **Mechanism** → **Injury outcome** → **Fatal or not**

Electrical exposure, 691 | Struck by object, 1,110 | Caught in or crushed, 2,179 | Fall to lower level, 923

$$\mathrm{RD}_X = P\big(F = 1 \mid \mathrm{do}(X = 1)\big) - P\big(F = 1 \mid \mathrm{do}(X = 0)\big)$$

X is a factor from the chain, 1 if the accident has it and 0 if not.

O is the injury outcome of the graph, such as caught in or crushed.

P(F = 1 | do(X = 1)) means the share of accidents that would be fatal if every accident had X.

F is fatality, 1 if someone died (OSHA's flag, or the extracted severity for Vertikal).

do(X = 1) means set X present in every accident, as in an experiment, instead of only observing it.

RD is the risk difference, so +0.16 means 16 percentage points more fatal accidents.

> Every adjustment also includes two background variables, the report source and the 2015 OSHA rule.

Pearl (2009), Causality, 2nd ed., Cambridge University Press

Determination of Conditional Independence

PC Algorithm, Stable Version

1. Start with every pair of variables linked.
2. Remove a link when the pair is independent, with conditioning sets of growing size.
3. Orient X to W to Y as a collider when W did not separate X and Y.
4. Orient the remaining links with the rules of Meek (1995).

causal-learn library, α = 0.01. With hundreds of tests, α acts as a sparsity setting.

G² Test, Worked Example

$$G^2 = 2 \sum_{x,y,z} N_{xyz} \ln \frac{N_{xyz}}{E_{xyz}}, \qquad E_{xyz} = \frac{N_{x+z}\, N_{+yz}}{N_{++z}}$$

N is the observed count, E the count expected under independence, and + sums over that index.

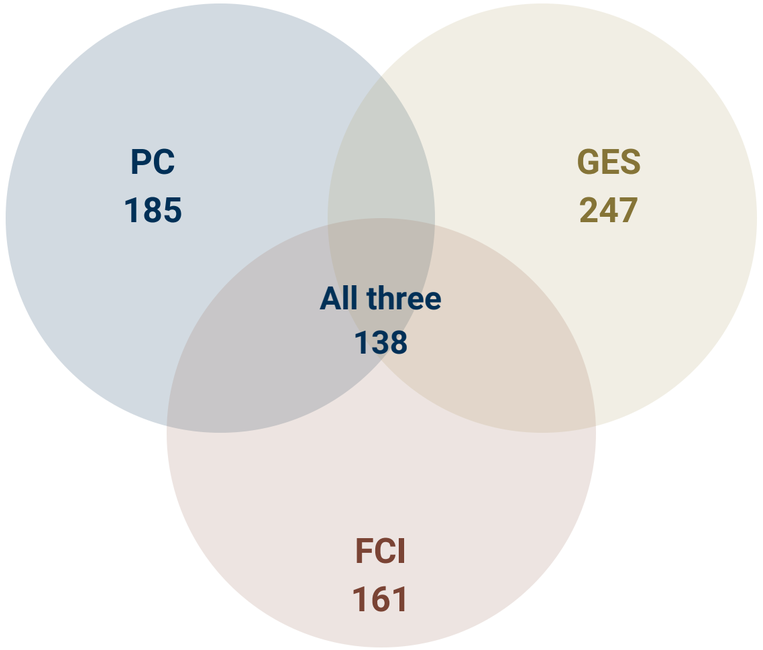
Observed (expected)	Not tipover	Tipover
Capacity not exceeded	4,418 (4,241.9)	990 (1,166.1)
Capacity exceeded	71 (247.1)	244 (67.9)

Marginal test: $G^2 = 482.2$, $p = 7.0 \times 10^{-107}$.
Within crane type: $G^2 = 347.4$, $p = 3.6 \times 10^{-76}$. The link stays.

> Assumptions: causal Markov, faithfulness, and causal sufficiency, the weakest here.

Spirtes and Glymour (1991), Social Science Computer Review 9(1) · Spirtes, Glymour, and Scheines (2001), Causation, Prediction, and Search, MIT Press · Meek (1995), Proc. UAI · Zheng et al. (2024), causal-learn, JMLR 25

Three Discovery Methods and Bootstrap Support



Links found by each method, summed over the four graphs. The union keeps all 280.

Graph	Union	By all 3	By 1	Stable
Electrical exposure	34	18	12	25
Struck by object	81	39	31	55
Caught in or crushed	78	38	27	49
Fall to lower level	87	43	35	57
Total	280	138	105	186

Stable means a bootstrap support of at least 0.9. Support is the share of 50 reruns on random 80% subsamples that contain the link.

All three methods agree on 49% of the links. Links found by one method are the weakest and are drawn dashed.

Spirtes, Glymour, and Scheines (2001), Causation, Prediction, and Search, MIT Press · Chickering (2002), JMLR 3 · Zheng et al. (2024), causal-learn, JMLR 25

Once we have the graph, how do we measure how strong a cause is?

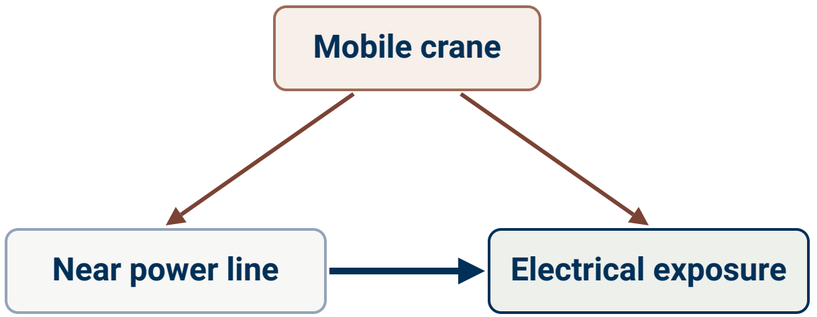
Compare Accidents That Are Alike Except for the Factor

$$P\big(Y \mid \mathrm{do}(X = x)\big) \;=\; \sum_{z} P(Y \mid X = x,\, Z = z)\; P(Z = z), \qquad \hat{p}_z \;=\; \frac{s_z + \bar{y}}{c_z + 1}$$

- **Y** is the outcome, here electrical exposure.
- **X** is the factor, here work near a power line, with x = 1 near a line.
- **Z** is the common causes found in the graph, here mobile or not, plus the state of the crane etc.
- **P(Y | X = x, Z = z)** is the share with the outcome inside one group, such as 92% for non-mobile cranes near a line.
- **P(Z = z)** is the weight of each group, its share of all accidents, 0.56 and 0.44.
- **p̂z** is the same group share with one extra accident at the overall share ȳ, so a tiny group cannot swing the result.

Effect = P(Y | do(X = 1)) − P(Y | do(X = 0)) = 0.56 × 0.90 + 0.44 × 0.85 = 0.88

Example: does work near a power line cause electrical exposure?



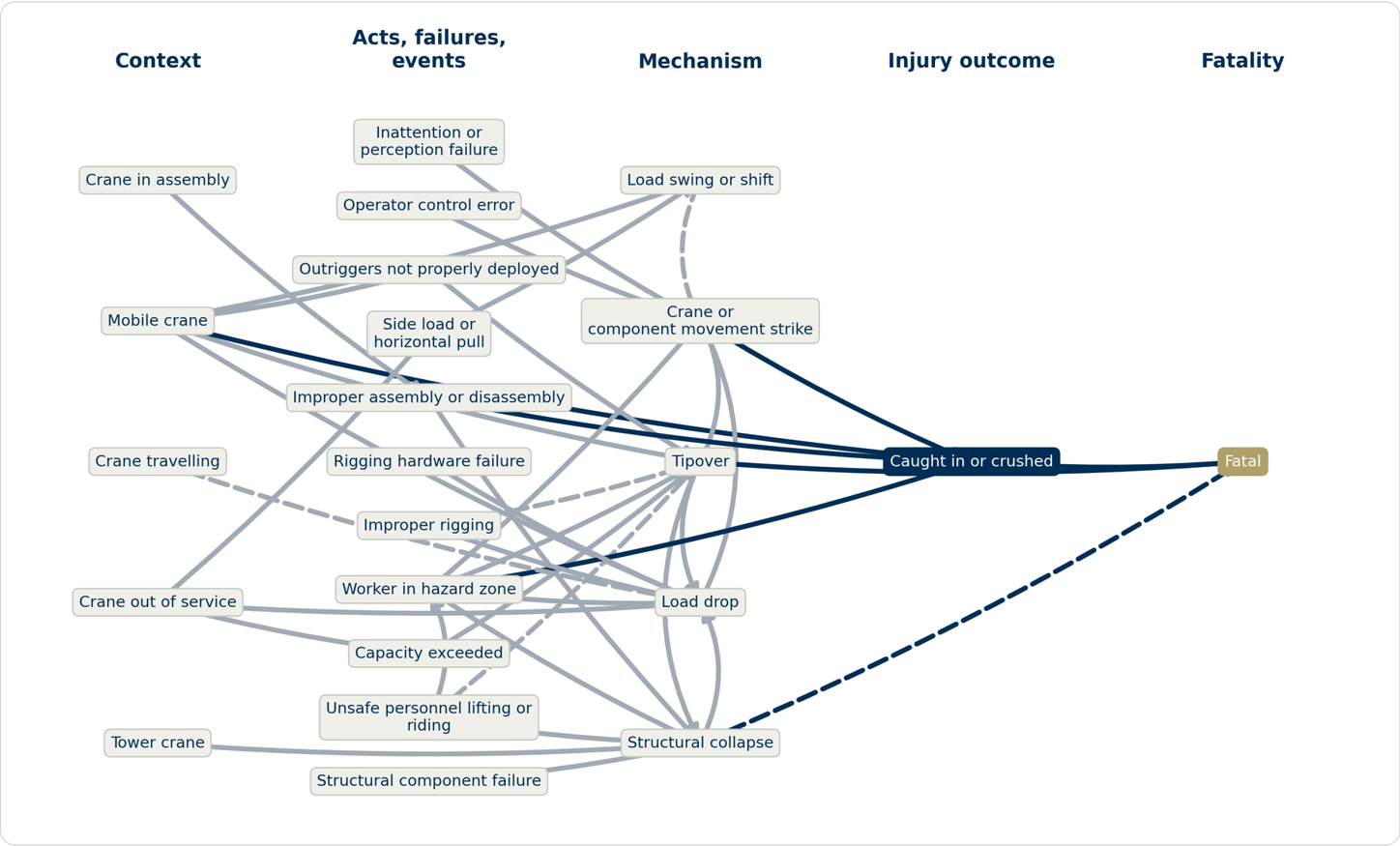
Group	Weight	Near a line	Otherwise	Difference
Not mobile	0.56	92%	1.4%	0.90
Mobile	0.44	86%	0.9%	0.85
Weighted				0.88

> Comparing like with like, the raw gap of 0.94 becomes 0.88 (95% interval 0.85 to 0.91).

Backdoor criterion, Pearl (2009), Causality, 2nd ed., Cambridge University Press

What does a learned graph look like?

Every Arrow Runs Forward, from Context to Fatality

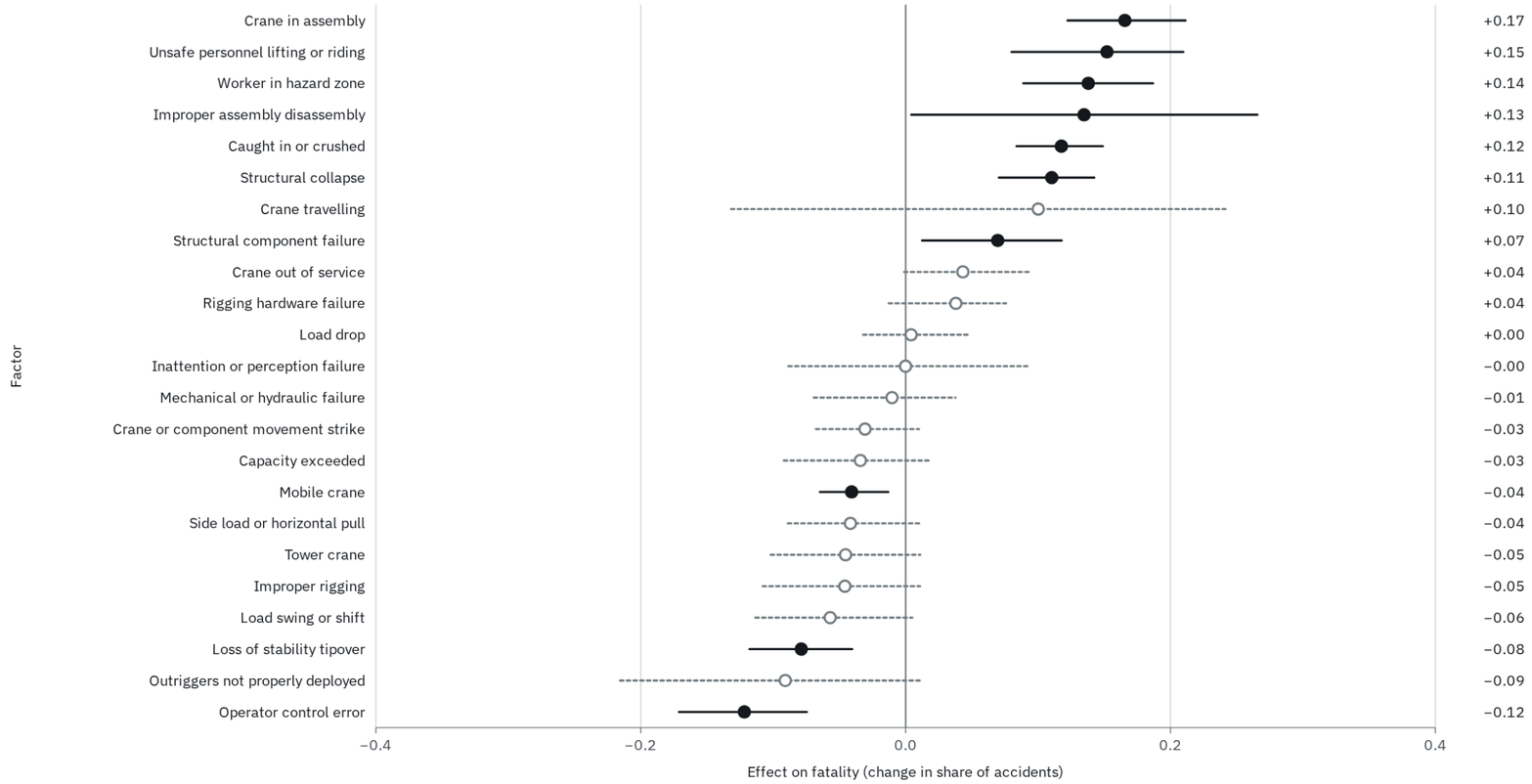


Caught in or Crushed

- The most common injury outcome, 2,179 accidents, 57% fatal
- Shown: the 40 links with bootstrap support of at least 0.95, out of 78 in the union
- Columns are the tiers, line width is support, dashed lines come from one method, dark lines enter the outcome or fatality

Within one graph, which factors move the chance of death, and how sure are we?

Effects on Fatality, Caught in or Crushed Graph



Each dot is an adjusted risk difference on fatality, with its 95% interval.

Filled dots and solid lines mark intervals that exclude zero.

Assembly work, lifting people, and the hazard zone raise the risk most. Tipover and operator error lower it.

Figure 13 of the technical note. Backdoor adjustment, Pearl (2009). The other three graphs are in the backup slides.

2. CAUSAL MODELING

What leads to each kind of injury?

Each Outcome Has Its Own Causes and Insights

1. **Electrical exposure comes almost entirely from work near power lines.** Power line proximity **+0.88**, power line contact **+0.53**, clearance not maintained **+0.35**. **Insight:** The risk is set before the lift, by where the crane and boom are placed.

2. **Struck by object comes from crane parts that break.** Wire rope or hoist line failure **+0.34**, structural component failure **+0.22**, rigging hardware failure **+0.20**. **Insight:** All three causes are equipment condition, so inspection covers them together.

3. **Caught in or crushed comes from people standing where the crane moves.** Worker in hazard zone **+0.25**, crane in assembly **+0.22**, crane or component movement strike **+0.12**. **Insight:** Assembly work carries this risk as well as the lift itself.

4. **Fall to lower level comes from using the crane to carry people.** Lifting or riding people on the crane **+0.58**, fall from crane or platform **+0.32**, inadequate safety management **+0.19**. **Insight:** The largest cause is a work practice, not an equipment failure.

Adjusted risk differences, all with 95% intervals that exclude zero. Insights come from observational data, not tested interventions.

Backdoor adjustment, Pearl (2009), Causality, 2nd ed., Cambridge University Press

What can a database of accident reports tell us?

Reports Show How Accidents End but Not How Often They Happen

All lifts
No records of safe lifts

→ Most lifts end safely

Accidents
Some are never reported

→ Severe ones are reported more often

Data
5,723 reported crane accidents

What the Reports Can Answer

Once an accident happens, does a factor make death more likely? A worker in the hazard zone raises the chance by about 0.14.

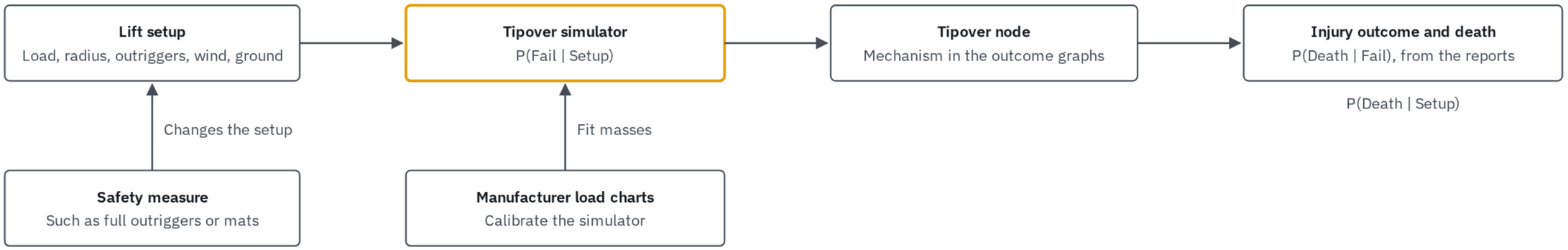
What They Cannot Answer

Does a factor make an accident more likely to happen? That needs the lifts that went well, which no report describes.

> Reports give how accidents end. How often they happen needs exposure data or physics, which is why the case study adds a tipover simulator.

Pearl (2009), Causality, 2nd ed., Cambridge University Press

Coupling the Graph with a Tipover Simulator



$$P(\text{Death} \mid \text{Setup}) = P(\text{Fail} \mid \text{Setup}) \times P(\text{Death} \mid \text{Fail})$$

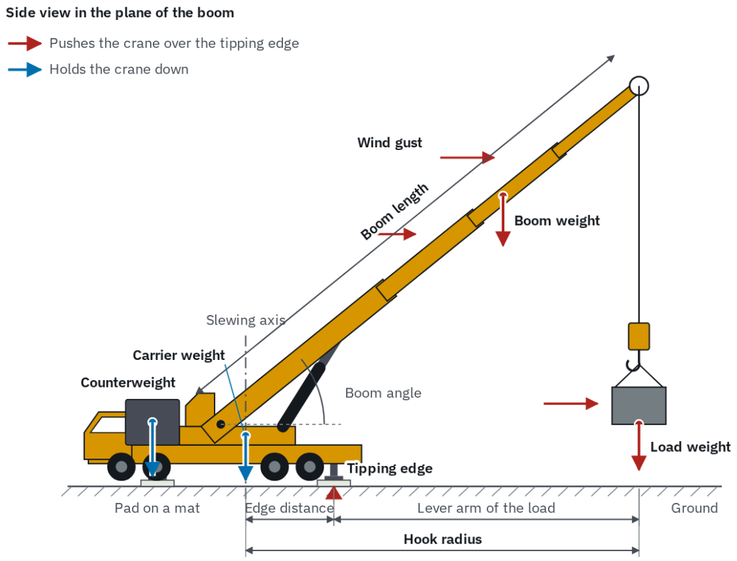
FROM THE SIMULATOR

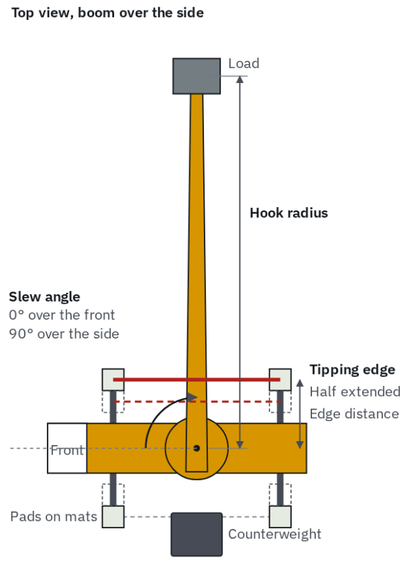
P(Fail | Setup) for any setup, including lifts that go well. Fail means the crane tips or the ground gives way.

FROM THE REPORTS

P(Death | Fail) = 28% over 1,234 tipovers (52% OSHA, 17% Vertikal)

Static Stability Model of a Mobile Crane





Moment Balance per Edge

$$U = \max_{\text{edges}} \frac{\sum M_{\text{overturning}}}{\sum M_{\text{restoring}}}$$

The crane tips when U ≥ 1 about any edge of the pad area. Each outrigger beam has its own extension.

Weights: payload, boom, counterweight, carrier. Lateral: wind and rope swing.

Pad load over mat area gives the ground pressure.

Started from a stability code shared by a collaborating group (M. Greenberg). We corrected it, extended it to three dimensions, and calibrated it.

Reference Crane and Rating Rules

Reference Crane



Parameter	Value
Crane	Liebherr LTM 1090-4.1, 90 t all-terrain
Outrigger base	8.5 m long, 7.0 m wide when fully extended
Minimum width	2.5 m, outriggers retracted
Boom	Telescopic, 11.1 to 50 m
Counterweight	21 t or 17 t, 4.5 m behind the slewing axis

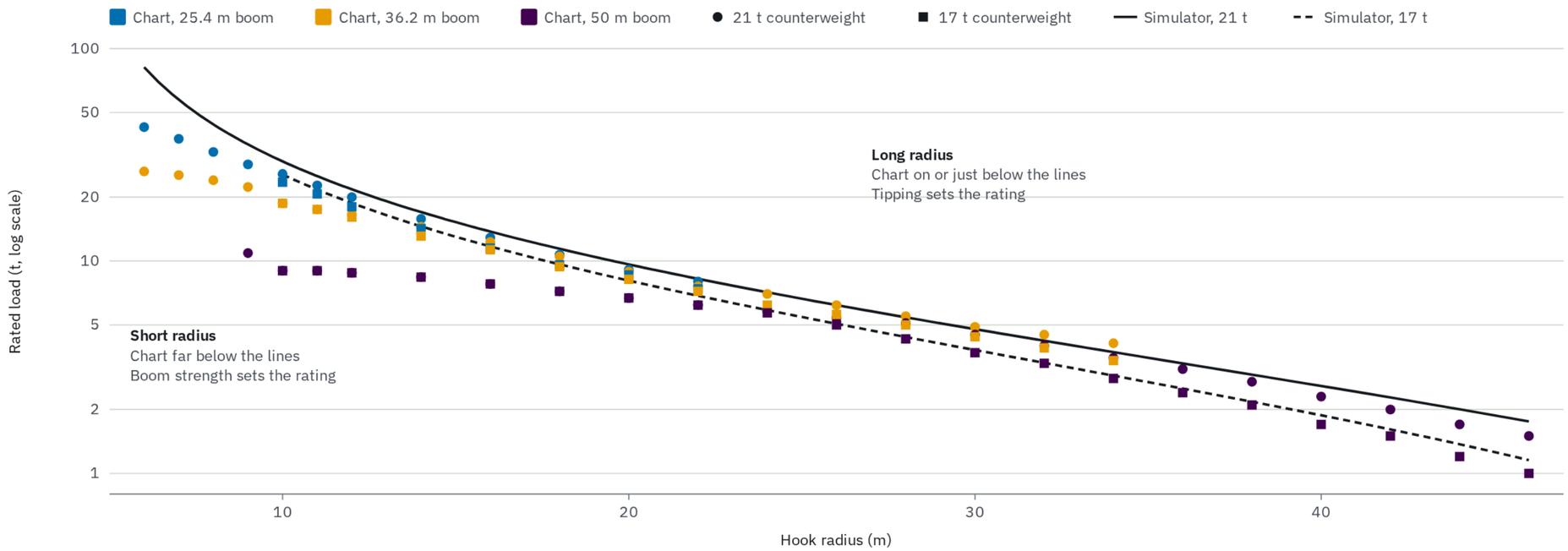
Calibrated Values and Rules

Parameter	Value
Carrier and superstructure	25.3 t, effective
Boom mass	5.5 t, effective
Boom center of gravity	68% of the boom length
Rating rule	EN 13000, stable under 1.25 × rated load plus 10% of boom weight
Wind	3 s gust at the boom tip (FEM 5.016)

> The fitted crane weighs 31 t without counterweight, against 48 t for the real crane. The lower effective masses absorb the wind, slope, and dynamic checks of EN 13000.

Does the simulator put the tipping limit where the manufacturer does?

The Fitted Limit Matches Charts the Fit Never Saw



Counterweight helps against tipping but not boom strength. Entries that drop when 4 t is removed are set by tipping, and these entries fit the three unknown masses.

> At short radius the boom strength sets the rating. Beyond about 10 to 24 m, depending on boom length, tipping sets it and the simulator matches the chart.

Monte Carlo Treatment of Input Uncertainty

Input	Distribution	Spread
Stated load	Lognormal	10% default, 25% if estimated, 5% if weighed
Hook radius	Normal, plus 2% boom deflection	0.3 m
Gust at boom tip	Normal multiplier	25%
Rope swing	Half-normal	2°
Ground bearing pressure	Lognormal	30%

4,000 to 20,000 samples per setup. The share where the crane tips or a pad sinks is P(Fail | Setup).

ONE INTERFACE FOR ANY SIMULATOR

```
simulate({
  "load_kg": 6000,
  "radius_m": 22, "boom_m": 29,
  "slew_deg": 90,
  "outrigger_ext": 0.5,
  "wind_ms": 12,
  "ground_kpa": 300,
  "mat_area_m2": 1.0})
```

Returns utilization, margin, p_tip, p_ground_failure, and the dominant inputs.

> A collapse, load drop, or power line model plugs into its own mechanism node through the same interface.

FEM 5.016 (2017), 3 s gust at the boom tip

Does the simulator agree with real accidents?

Simulator and Report Agree on 77% of Clear Verdicts

Simulator says	Reports	Report agrees	Agreement
Stable, p below 0.1	73	58 did not tip	79%
Tips, p above 0.9	9	5 tipped	56%
Both	82	63	77%

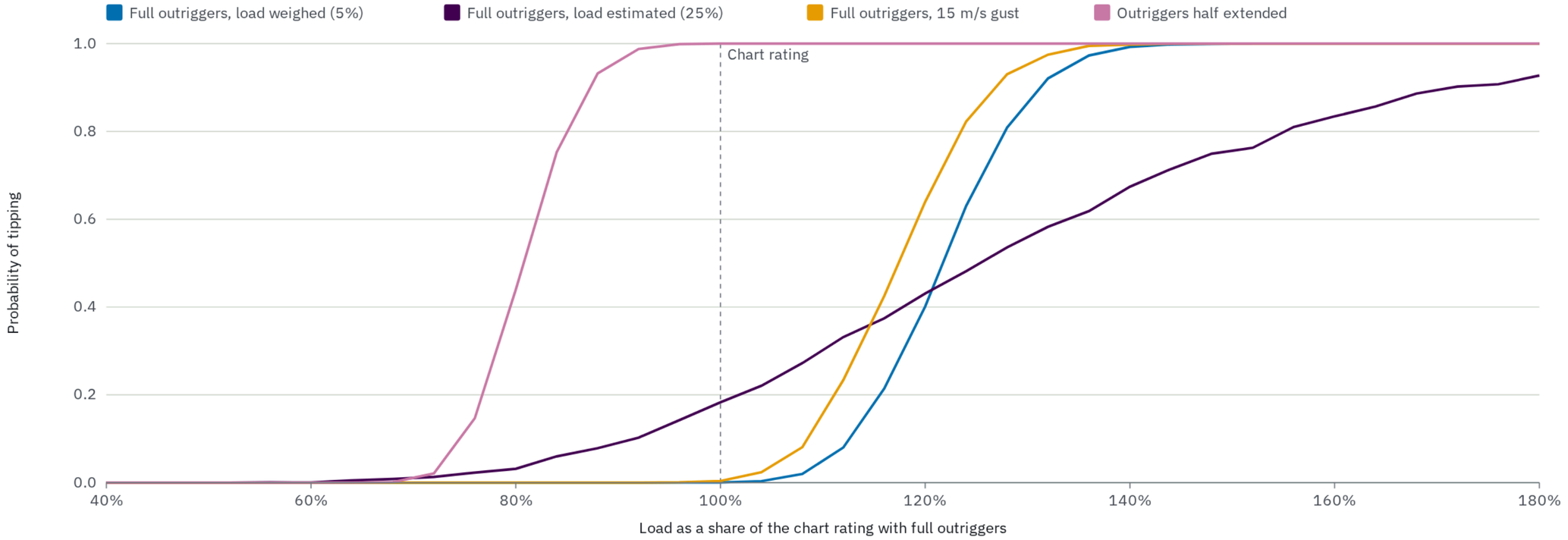
Left out: 23 reports where the simulator is uncertain (p from 0.1 to 0.9). 13 of them tipped.

> Most disagreements have a reason: 8 of the 15 "stable" lifts that tipped also name a ground failure, an outrigger error, or wind, which a load check cannot see.

3. TIPOVER SIMULATION CASE STUDY 1

How much margin does a chart rating leave, and what erodes it?

Half-Extended Outriggers Erode the Margin Most



A 15 m/s gust shifts the curve left. Of the inputs varied here, the outrigger setup has the largest effect.

Before a lift: which safety measure helps most for this planned lift?

Extending the Outriggers Cuts the Chance of Death Most

The plan (one synthetic lift):
5.8 t at 22 m radius on a 29 m boom, over the side; outriggers half extended, load weight estimated (±25%), 12 m/s gusts, 1 m² pads on 300 kPa gravel.
Each row changes one thing in the plan.
Chance of a death = chance of either failure × chance of a fatal accident among all accidents.

Change to the plan	Load Percentage	Crane tips	Ground gives way	Either failure	Chance of a death	Death risk vs. plan
None, as planned	89%	28%	40%	51%	14.3%	baseline
Extend the outriggers fully	66%	1%	22%	23%	6.4%	−55%
Use 3 m² timber mats	89%	28%	0%	28%	7.9%	−45%
Load uncertainty (±5% instead of ±25%)	89%	3%	41%	42%	11.9%	−17%
Wait for gusts under 9 m/s	89%	27%	39%	50%	14.1%	−1%
All four changes together	66%	0%	0%	0%	0.0%	−100%

> Every change lowers the risk, but by very different amounts. Outriggers and mats each cut it about in half; the gust limit barely moves it for this lift.

3. TIPOVER SIMULATION CASE STUDY 1

After an accident: was the outrigger error necessary for OSHA record 3521?

Extended Outriggers Would Almost Surely Have Prevented It

A mobile crane tipped while lifting an 862 kg panel with its outriggers not properly set. The report states the chart rating for both setups. It gives no ground data, so here Fail is the tipover.

Quantity	As set	Fully extended
Chart rating	227 kg	2,313 kg
Load to rating	3.8	0.37
P(Fail), simulator	97%	0%

PROBABILITY OF NECESSITY

$$\mathrm{PN} = 1 - \frac{P(\mathrm{Fail} \mid \mathrm{Extended})}{P(\mathrm{Fail} \mid \mathrm{As\ set})}$$

PN = 1.00 under monotonicity (Tian and Pearl, 2000)

> The graph links outrigger errors to tipover across many accidents. The simulator turns that link into a statement about one accident.

Tian and Pearl (2000), Annals of Mathematics and Artificial Intelligence 28 · OSHA, Accident Investigation Search (IMIS)

Which crane should we calibrate the simulator on?

We Calibrate on the Crane of a Fatal Tipover, a Terex T775



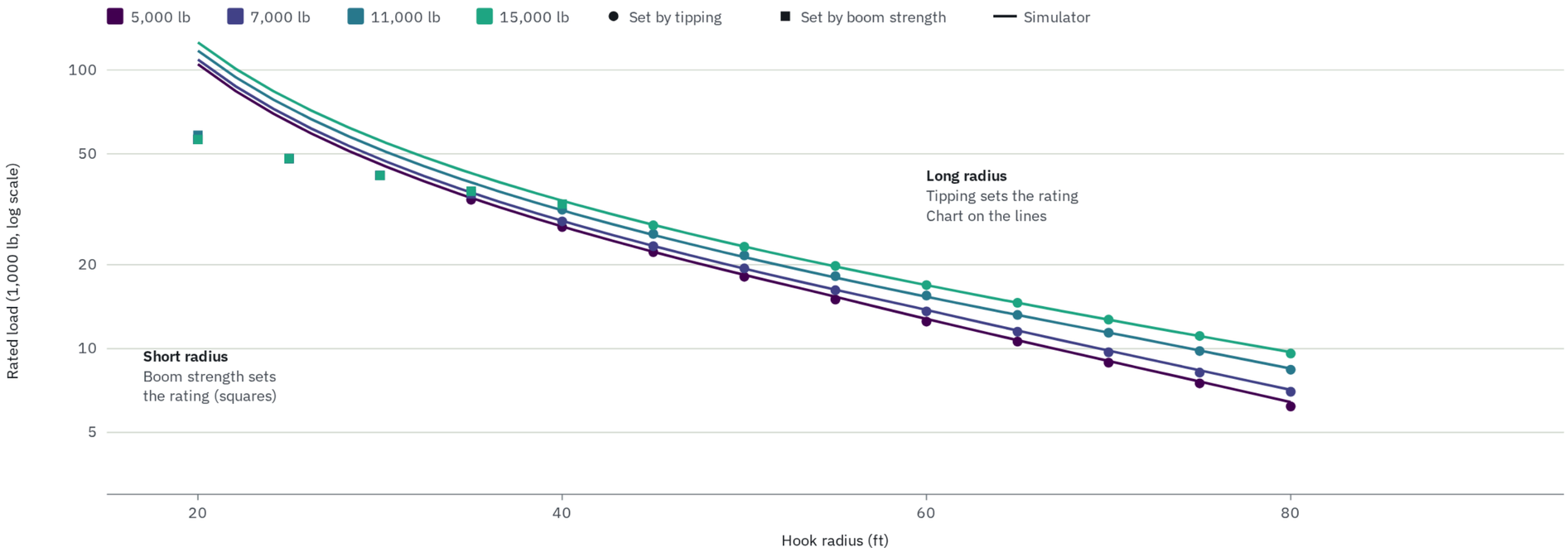
Parameter	Value
Crane	Terex T775, 75 US ton truck crane
Outrigger spread	26 ft (7.9 m) at full extension
Boom	40 to 126 ft, four sections
Counterweight	5,000, 7,000, 11,000, or 15,000 lb
Rating rule	85% of the tipping load (ASME B30.5)
Weight	36.0 t without counterweight, from the axle loads

Terex Cranes (2005), T775 lifting capacity and specifications · OSHA, Accident Investigation Search (IMIS)

3. TIPOVER SIMULATION CASE STUDY 2

Does the simulator put the tipping limit where the manufacturer does?

The Fit Also Predicts a Counterweight It Never Saw



Circles are ratings set by tipping, and squares are ratings set by boom strength. Lines are the simulator.

> The mean error is 1.5%, and 1.7% on the 7,000 lb chart of the accident when we leave it out of the fit. The fitted crane weighs 39.5 t, against 36.0 t on the spec sheet.

Terex Cranes (2005), T775 lifting capacity spec sheet · ASME B30.5, Mobile and Locomotive Cranes · SAE J765a

Can the simulator predict a chart rating it was never shown?

The Simulator Predicts the Accident's Rating Within 11%

Accident setup, OSHA record 3358 (June 2006, fatal)

- Terex T775 lifting a 44,775 lb cable reel, **46,383 lb** with the hook block
- 7,000 lb counterweight, 54.8 ft boom at 73°, radius 15 ft 9 in
- Load-side outriggers retracted

Out-of-sample test

- Chart rating quoted in the report for this setup: **14,000 lb**, so the reel was more than three times the rating
- Simulator prediction for the same setup: **15,519 lb**, 11% above the chart
- The gap rests on one assumption, the retracted pad spacing: 8 ft (the width of the truck) gives 11%, 7.5 ft gives an exact match

> The calibrated simulator holds up on a setup it never saw, so we use it to replay this accident.

OSHA, Accident Investigation Search (IMIS), record 3358 · Terex Cranes (2005), T775 specifications

Which single change before the lift would have prevented the tipover?

Outriggers or the Rated Load Would Have Prevented It

Each row changes one input. The 21.0 t hook load, the 16.7 m boom, and the slew over the side stay the same. The tipping radius is the hook radius beyond which the crane tips.

Setup	Tipping radius	P(tip), 3.45 m	P(tip), 4.80 m
As set, load-side beams retracted	3.67 m	74%	100%
All outriggers fully extended	10.49 m	0%	0%
All outriggers half extended	7.08 m	0%	0%
As set, with 15,000 lb of counterweight	4.34 m	14%	100%
As set, with the load at its 14,000 lb rating	6.97 m	0%	0%

> More counterweight does not help at the reported radius. Extending the load-side beams by 17% of their travel would have held the reel.

OSHA, Accident Investigation Search (IMIS), record 3358 · Terex Cranes (2005), T775 lifting capacity

Was each error in OSHA record 3358 necessary for the tipover?

Either Fix Alone Would Have Prevented the Tipover

The lift had two errors. The load-side beams were retracted, and the reel weighed 3.2 times the rating. The report gives no ground data, so Fail here means the tipover.

Setup	P(Fail)	PN, 3.45 m	PN, 4.80 m
As set	74% to 100%		
Outriggers fully extended	0%	1.00	1.00
Load at the 14,000 lb rating	0%	1.00	1.00
15,000 lb counterweight (for comparison)	14% to 100%	0.81 to 1.00	0.00

PROBABILITY OF NECESSITY

$$\mathrm{PN} = 1 - \frac{P(\mathrm{Fail} \mid \mathrm{Extended})}{P(\mathrm{Fail} \mid \mathrm{As\ set})}$$

PN = 1.00 for both errors under monotonicity (Tian and Pearl, 2000)

> Every mass in the simulator comes from the chart of the crane that tipped. The link in the graph from outrigger errors to tipover holds for this accident.

Is a tipover simulator enough to estimate the chance of death on a site?

Tipover Is Common but Least Fatal, 12% of Fatal Accidents

Primary mechanism	Accidents	Fatal share	Share of fatal accidents	Physics model needed	Status
Load drop	1,207	51%	22%	Rigging, hoist line, and load release	-
Movement strike	1,102	48%	19%	Slew and travel kinematics with worker positions	-
Power line contact	703	68%	17%	Boom reach and minimum approach distance	-
Structural collapse	777	60%	17%	Boom and structure capacity	-
Tipover	1,202	28%	12%	Calibrated stability model	Built, this talk
Fall from crane or platform	413	65%	10%	Personnel platform and fall protection	-

Limitations

1. Every record is an accident, so the graphs give no accident rates.
2. The simulator is static and rigid, and each accident needs its own crane's chart.
3. P(Death | Fail) from reports is closer to an upper bound.

Next Steps

1. Structural collapse simulator, in development.
2. Load drop and power line clearance models, about 40% of deaths.
3. Weather records for wind accidents, and exposure data for rates.
4. Contact the company and obtain its 701 forensic accident cases.
5. Draft the paper manuscript.

5,723 crane accidents, 2,793 fatal, mechanisms with at least 100 accidents · OSHA 29 CFR 1926 Subpart CC · Wiethorn (2018)

Causal Discovery and Inference

Chickering, D.M. (2002). Optimal Structure Identification With Greedy Search. Journal of Machine Learning Research 3, 507–554.

Colombo, D., Maathuis, M.H. (2014). Order-Independent Constraint-Based Causal Structure Learning. Journal of Machine Learning Research 15.

Meek, C. (1995). Causal Inference and Causal Explanation with Background Knowledge. Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence.

Pearl, J. (2009). Causality: Models, Reasoning, and Inference, second edition. Cambridge University Press.

Ramsey, J., Zhang, J., Spirtes, P. (2006). Adjacency-Faithfulness and Conservative Causal Inference. Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence.

Spirtes, P., Glymour, C. (1991). An Algorithm for Fast Recovery of Sparse Causal Graphs. Social Science Computer Review 9 (1), 62–72.

Spirtes, P., Glymour, C., Scheines, R. (2001). Causation, Prediction, and Search, second edition. The MIT Press.

Tian, J., Pearl, J. (2000). Probabilities of causation: Bounds and identification. Annals of Mathematics and Artificial Intelligence 28 (1–4), 287–313.

Vashishtha, A., Reddy, A.G., Kumar, A., Bachu, S., Balasubramanian, V.N., Sharma, A. (2025). Causal Order: The Key to Leveraging Imperfect Experts in Causal Inference. ICLR 2025.

Zheng, Y., Huang, B., Chen, W., Ramsey, J., Gong, M., Cai, R., Shimizu, S., Spirtes, P., Zhang, K. (2024). Causal-learn: Causal Discovery in Python. Journal of Machine Learning Research 25 (60), 1–8.

Crane Safety, Standards, Text Models, and Data

Aneziris, O.N., Papazoglou, I.A., et al. (2008). Towards risk assessment for crane activities. Safety Science 46 (6), 872–884.

American Society of Mechanical Engineers. ASME B30.5, Mobile and Locomotive Cranes.

Blei, D.M., Ng, A.Y., Jordan, M.I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research 3, 993–1022.

Bureau of Labor Statistics (2012). Occupational Injury and Illness Classification Manual, Version 2.01.

European Committee for Standardization. EN 13000, Cranes, Mobile cranes. European Materials Handling Federation (2017). FEM 5.016.

McCann, M., Gittleman, J., Watters, M. (2009). Crane-Related Deaths in Construction and Recommendations for Their Prevention. CPWR.

Raviv, G., Fishbain, B., Shapira, A. (2017). Analyzing risk factors in crane-related near-miss and accident reports. Safety Science 91, 192–205.

Shapira, A., Lyachin, B. (2009). Identification and Analysis of Factors Affecting Safety on Construction Sites with Tower Cranes. J. Constr. Eng. Manage. 135 (1), 24–33.

Shappell, S.A., Wiegmann, D.A. (2000). The Human Factors Analysis and Classification System, HFACS. Report DOT/FAA/AM-00/7, FAA.

Wiethorn, J.D. (2018). An Analytical Study of Critical Factors of Lift Planning to Improve Crane Safety Based on Forensic Case Studies of Crane Accidents. PhD dissertation, The University of Texas at Austin.

Occupational Safety and Health Administration. 29 CFR 1926 Subpart CC, Cranes and Derricks in Construction. Accident Investigation Search, IMIS.

Liebherr. LTM 1090-4.1 and LTM 1090-4.2 technical data and load charts. Terex Cranes (2005). T775 lifting capacity and specifications.

Vertikal.net, news for the crane and access industry.

6. SUMMARY

Three Questions, Three Answers

RQ1. Can a language model turn reports into reliable causal chains? **Yes, with an enforced schema.**

1. **Every output fits the schema.** Each report maps to 55 factor types, and every link quotes the sentence behind it.
2. **Agreement with human labels.** Kappa is 0.78 for wind and 0.52 to 0.65 for crane type, where 1 means perfect agreement.

RQ2. Which factors cause each outcome, and which make it fatal? **They can be shown with a causal graph.**

1. **The type of injury matters most.** Electrical exposure is 71% fatal. Each injury type still adds 0.09 to 0.21 after adjustment.
2. **Two factors raise the risk in all four graphs.** Work during assembly adds 0.17, and a worker in the hazard zone adds 0.13 to 0.15.

RQ3. How likely is a lift to fail, and which measure helps most? **A calibrated simulator supplies the failure probability.**

1. **The simulator matches the manufacturer's charts.** Case 1 on a Liebherr, case 2 on the Terex T775 of a fatal tipover, 1.5% error.
2. **It finds what was necessary in one accident.** In OSHA 3358, full outriggers or a load at its rating each prevent the tipover.
3. **It ranks safety measures.** In the example lift, extending the outriggers cuts the chance of death from 14.3% to 6.4%.
4. **Tipover is only 12% of fatal accidents.** Collapse, load drop, and power line models come next.

Thank you. Questions?

Worked G² Test for Capacity Exceeded and Tipover

Cell	Observed N	Expected E	2 N ln(N/E)
Not exceeded, not tipover	4,418	4,241.9	+359.4
Not exceeded, tipover	990	1,166.1	−324.1
Exceeded, not tipover	71	247.1	−177.1
Exceeded, tipover	244	67.9	+624.1

DECISION

p is far below α = 0.01 in both tests, so PC keeps the link, even within crane type.

$$(1)\quad E_{ij} = \frac{N_{i+}\,N_{+j}}{N}, \qquad E_{\text{exceeded, tipover}} = \frac{315 \times 1234}{5723} = 67.9$$

$$(2)\quad G^2 = \sum_{i,j} 2\,N_{ij} \ln \frac{N_{ij}}{E_{ij}} = 359.4 - 324.1 - 177.1 + 624.1 = 482.2$$

$$(3)\quad \mathrm{df} = (2-1)(2-1) = 1, \qquad p = P\left(\chi^2_1 \geq 482.2\right) = 7.0 \times 10^{-107}$$

$$(4)\quad G^2_{|Z} = G^2_{\text{other}} + G^2_{\text{mobile}} = 100.3 + 247.1 = 347.4, \qquad \mathrm{df} = 2, \quad p = 3.6 \times 10^{-76}$$

Step 4 counts (exceeded and tipover, exceeded only, tipover only, neither): other cranes 38, 33, 230, 2,886, and mobile cranes 206, 38, 760, 1,532 · Spirtes, Glymour, and Scheines (2001)

The Two Phases of the PC Algorithm

Phase 1, Skeleton

1. Start from the complete undirected graph.
2. For $\ell = 0, 1, 2$, and so on, test each linked pair X and Y given every set S of ℓ other neighbours of X.
3. If a test accepts independence, delete the link and store S as the separating set of X and Y.
4. Stop when no node has more than ℓ neighbours left.

Cost grows with the largest node degree, so PC is fast on sparse graphs.

Phase 2, Orientation

1. Colliders: for X and Y not linked, both linked to W, orient X into W and Y into W when W is not in their separating set.
2. Background knowledge: forbid any arrow from a later tier to an earlier tier.
3. Meek rules R1 to R4: orient further links whenever another direction would create a new collider or a cycle.

The output is a CPDAG. Links left undirected are excluded from effect estimation.

Our run: PC-stable, G^2 test, $\alpha = 0.01$, seven tiers as background knowledge, causal-learn.

Spirtes and Glymour (1991), Social Science Computer Review 9(1) · Meek (1995), Proc. UAI · Zheng et al. (2024), causal-learn, JMLR 25

PC and Its Variants, and What We Used

Variant	What changes	Main benefit	Used here
Original PC	Deletes a link as soon as a test accepts independence	Fast	No
PC-stable	Fixes the neighbour sets at the start of each level and deletes links only at its end	Skeleton does not depend on variable order	Yes, main method
Conservative PC	Orients a collider only if W is in none of the separating sets found, and flags ambiguous triples	Fewer false colliders	No
Majority-rule PC	Orients a collider by the majority of separating sets	Order-independent orientations	No
FCI	Allows hidden common causes and selection, and returns a PAG with circle marks	No causal sufficiency needed	Yes, as a check

GES, the score-based method in our union, searches over equivalence classes with the BDeu score instead of testing independence.

Spirtes and Glymour (1991) · Colombo and Maathuis (2014), Order-Independent Constraint-Based Causal Structure Learning, JMLR 15 · Ramsey, Zhang, and Spirtes (2006), Proc. UAI · Spirtes et al. (2001) · Chickering (2002), JMLR 3

Effects on Fatality, Struck by Object Graph

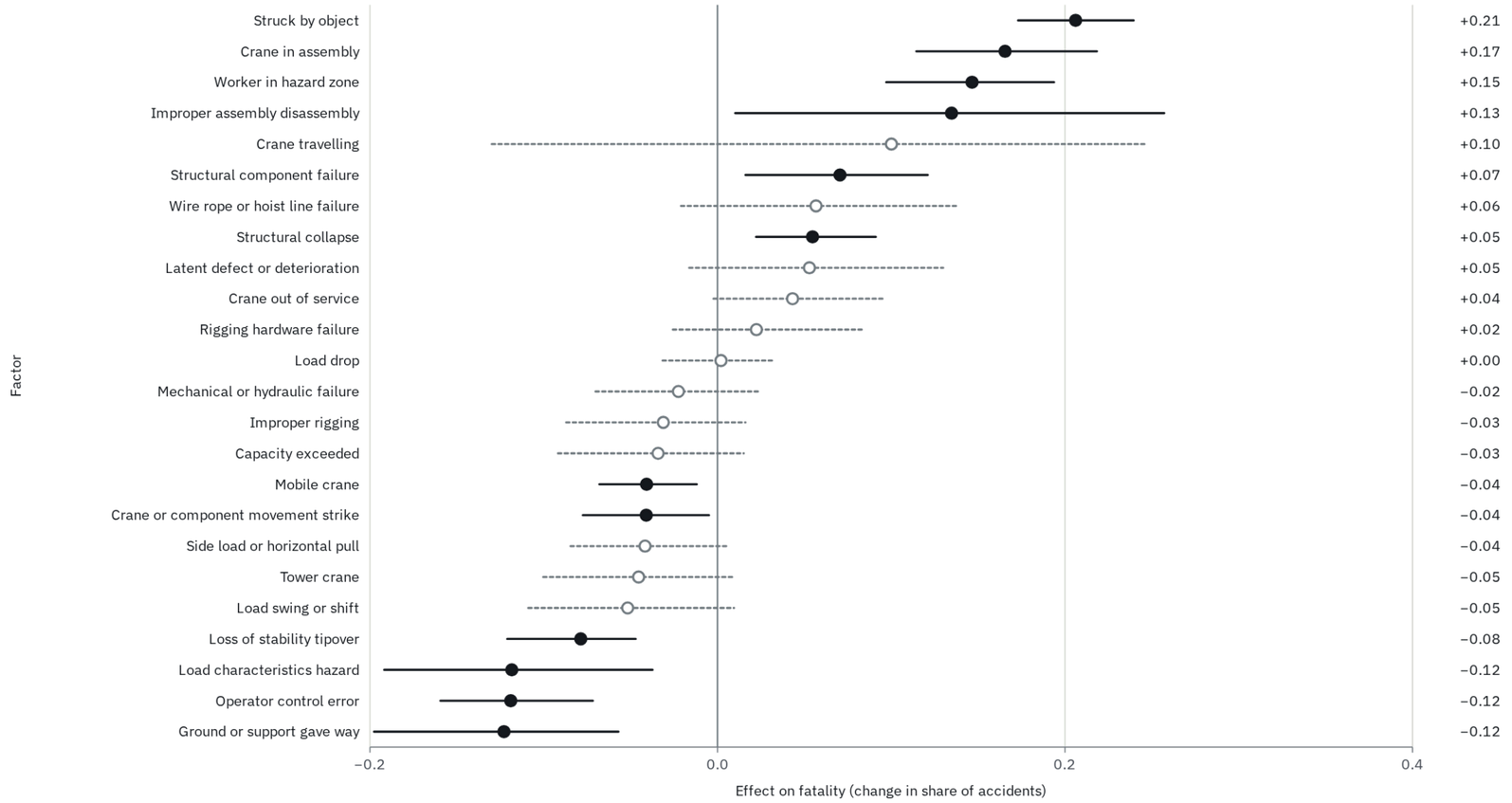


Figure 13 of the technical note. Filled dots and solid lines mark 95% intervals that exclude zero. Backdoor adjustment, Pearl (2009).

BACKUP · EFFECTS ON FATALITY

Effects on Fatality, Electrical Exposure Graph

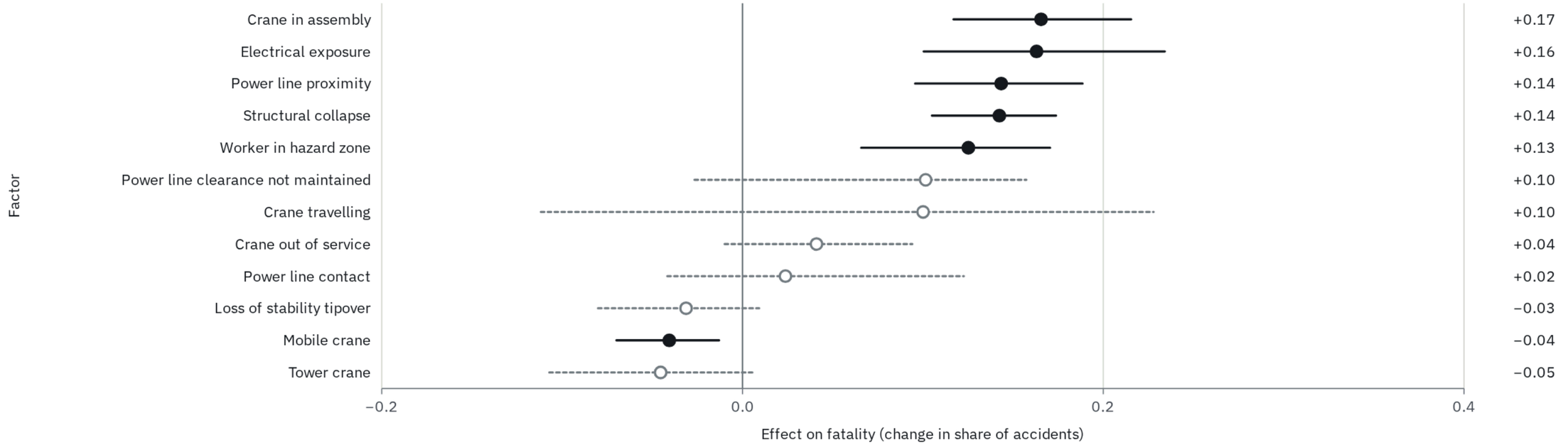


Figure 13 of the technical note. Filled dots and solid lines mark 95% intervals that exclude zero. Backdoor adjustment, Pearl (2009).

Effects on Fatality, Fall to Lower Level Graph

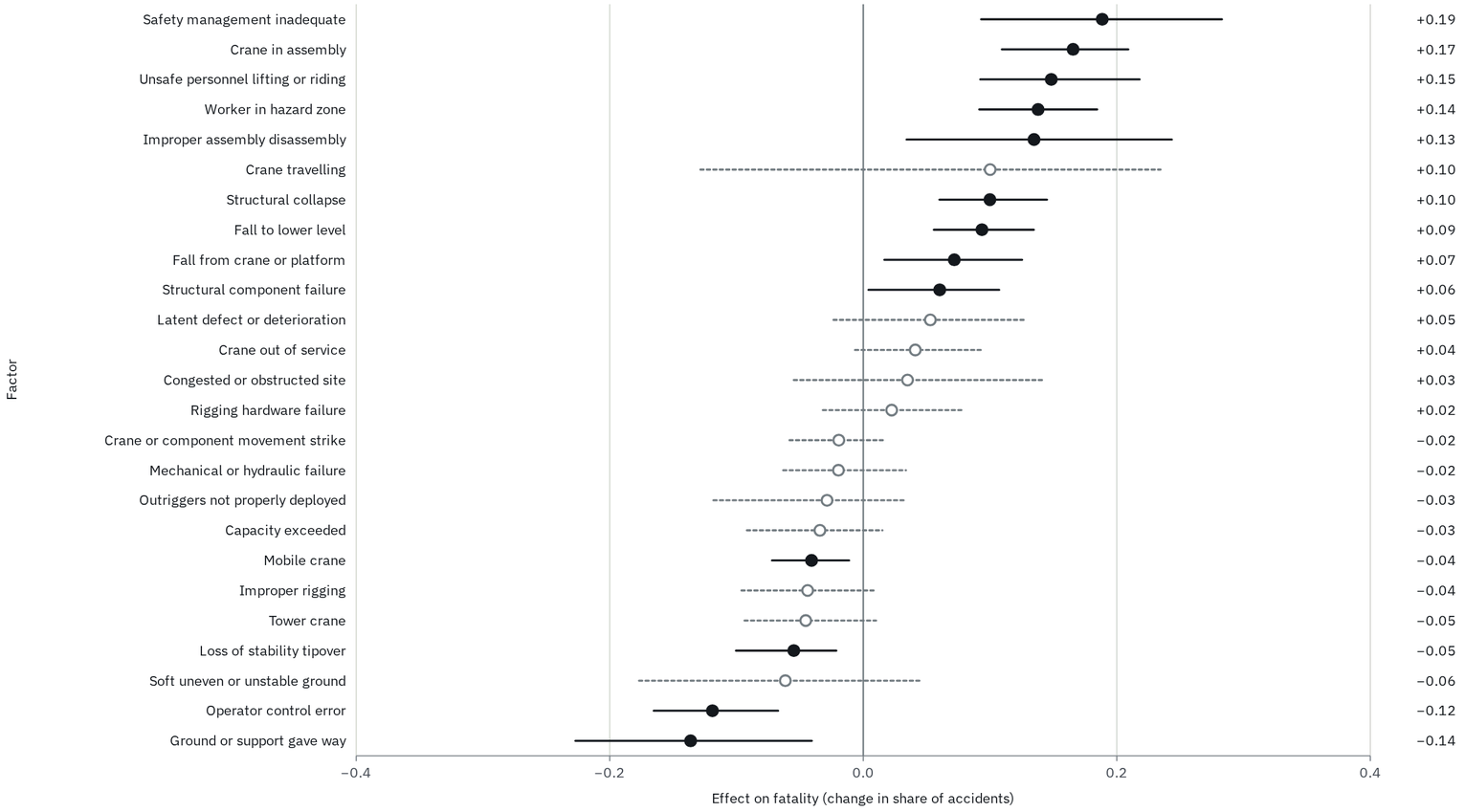


Figure 13 of the technical note. Filled dots and solid lines mark 95% intervals that exclude zero. Backdoor adjustment, Pearl (2009).

METHOD 1 · FACTOR TYPES, 1 OF 2

Organization, Preconditions, Environment, Equipment

Organization (5)

- Lift planning inadequate
- Supervision inadequate
- Inspection or maintenance inadequate
- Safety management inadequate
- Design or manufacturing defect

Preconditions (3)

- Operator unqualified or untrained
- Fatigue or impairment
- Limited visibility or blind lift

Environment (8)

- High wind condition
- Wind gust or wind load event
- Other adverse weather
- Soft, uneven, or unstable ground
- Ground or support gave way
- Power line proximity
- Congested or obstructed site
- Load characteristics hazard

Equipment (7)

- Wire rope or hoist line failure
- Rigging hardware failure
- Structural component failure
- Mechanical or hydraulic failure
- Safety device missing or malfunctioning
- Latent defect or deterioration
- Equipment unsuitable for task

Shappell and Wiegmann (2000), HFACS · Shapira and Lyachin (2009), J. Constr. Eng. Manage. 135 · Raviv et al. (2017), Safety Science 91 · OSHA 29 CFR 1926 Subpart CC · FEM 5.016 (2017)

METHOD 1 · FACTOR TYPES, 2 OF 2

Unsafe Acts, Mechanisms, and Outcomes

Unsafe Acts (17)

- Capacity exceeded
- Load weight or chart misjudged
- Outriggers not properly deployed
- Improper crane setup or positioning
- Side load or horizontal pull
- Improper rigging
- Operator control error
- Inattention or perception failure
- Signal or communication failure
- Safety device bypassed
- Power line clearance not maintained
- Worker in hazard zone
- Unsafe personnel lifting or riding
- Improper assembly or disassembly
- Out of service, not secured
- Decision to continue in adverse conditions
- Unsafe travel or pick and carry

Mechanism (9)

- Loss of stability, tipover
- Structural collapse
- Load drop
- Load swing or shift
- Power line contact
- Two-blocking
- Crane or component movement strike
- Collision
- Fall from crane or platform

Outcome (6)

- Electrical exposure
- Struck by object
- Caught in or crushed
- Fall to lower level
- Struck against or other injury
- Property or equipment damage only

OSHA 29 CFR 1926 Subpart CC · McCann, Gittleman, and Watters (2009), CPWR · Aneziris et al. (2008), Safety Science 46(6) · BLS (2012), OIICS Manual 2.01

Record-Level Vocabularies

Equipment Class (14)

- Mobile telescopic, truck or all terrain
- Mobile rough terrain
- Crawler lattice
- Boom truck or knuckle boom
- Tower or self-erecting
- Overhead, gantry, or bridge
- Floating or barge
- Rail mounted
- Derrick
- Other crane
- Unknown crane type
- Aerial work platform, not a crane
- Non-crane lifting equipment
- Unknown

The first eleven form the crane population. Source codes follow BLS OIICS.

Crane State (7)

- In service, lifting
- In service, no load
- Out of service
- Assembly or disassembly
- Travelling
- Maintenance
- Unknown

Severity (7)

- Fatal
- Severe injury
- Minor injury
- Severe damage
- Slight damage
- No injury or damage
- Unknown

Record Type (2)

- Accident report
- Not an accident report

Operation Phases and Physical Drivers

Operation Phase (14)

- Transport or road travel
- Setup or positioning
- Assembly, disassembly, or erection
- Rigging or hooking
- Hoisting
- Swinging or slewing
- Booming or telescoping
- Travelling with load
- Lowering or landing
- Unrigging
- Idle or out of service
- Maintenance or inspection
- Not applicable
- Unknown

Adapted from the activity field of Raviv et al. (2017).

Drivers (8)

- Load weight
- Rated capacity at configuration
- Crane nominal capacity
- Radius
- Boom length
- Boom angle
- Wind speed
- Power line voltage

Each a value and unit as written, plus the crane make and model

Outrigger State (5)

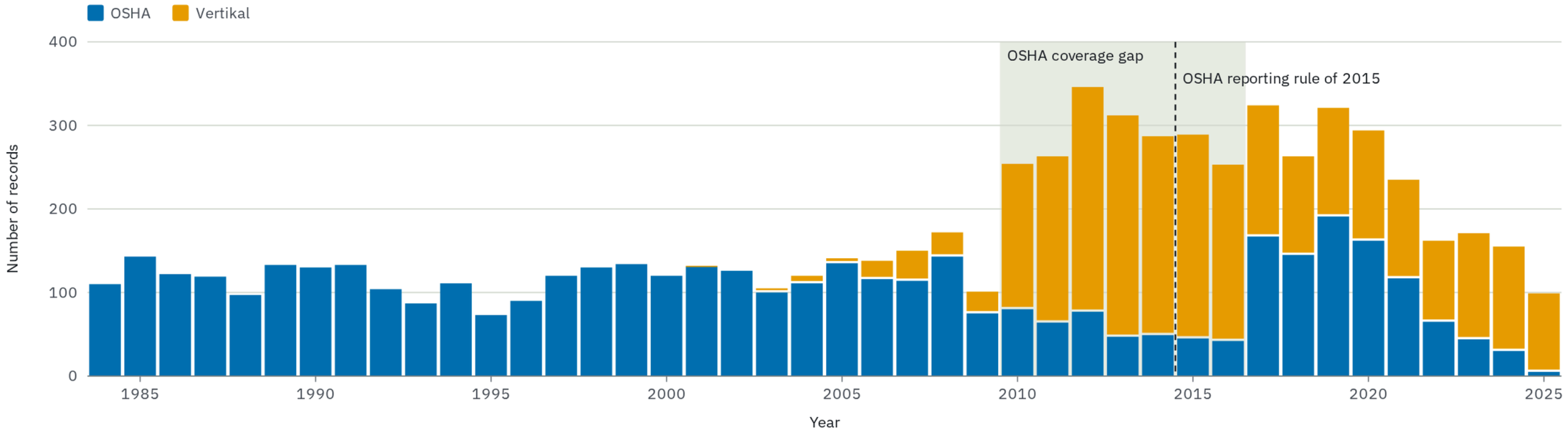
- Fully extended
- Partially extended
- Not deployed
- Not applicable
- Unknown

Wind Kind (4)

- Gust
- Sustained
- Wind mentioned, kind unclear
- No wind mentioned

DATA

Yearly Counts Do Not Show Safety Trends

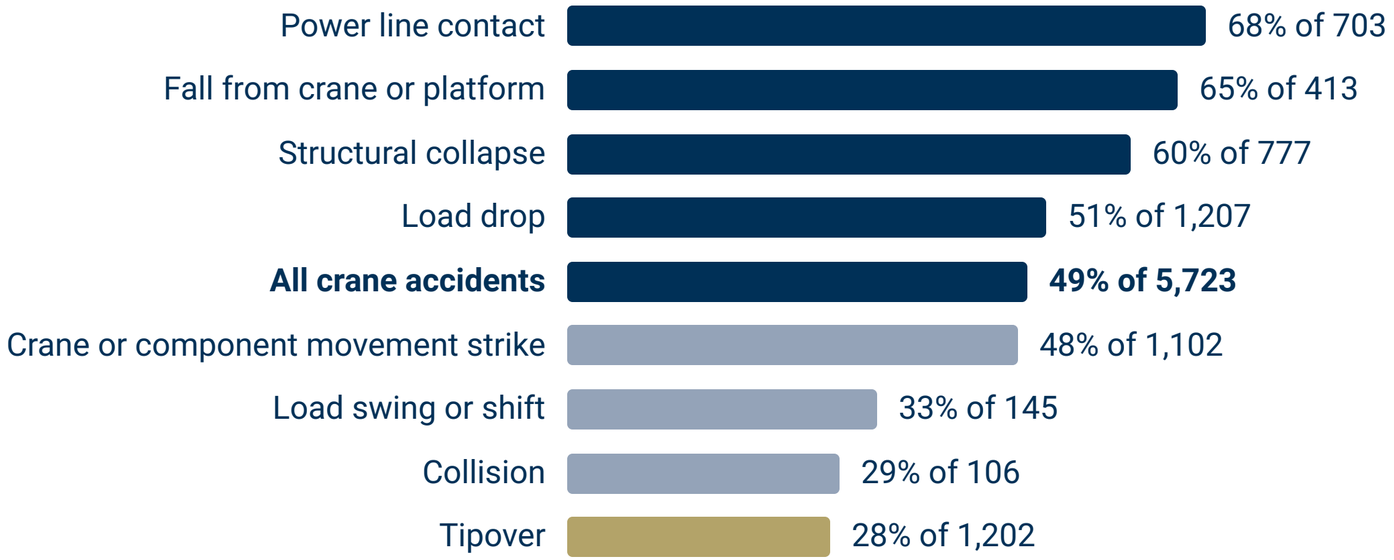


OSHA has a collection gap from 2010 to 2016, and a 2015 rule added hospital stays and amputations. The models adjust for source and the 2015 rule.

OSHA, Accident Investigation Search (IMIS)

RESULTS · FATAL ACCIDENTS

Fatal Share by Primary Mechanism



TIPOVER

Second most common and least fatal. 47% of tipovers damage only property.

DEADLIEST

Power line contact, falls from the crane, and structural collapse

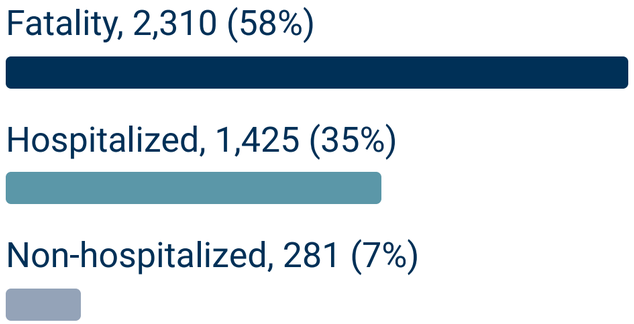
Two-blocking (10 accidents) and records without a mechanism (58) are left out.

RESULTS

Predicting the Severity of Injury

Target, OSHA Degree of Injury

4,016 OSHA crane accidents, most severely injured victim. The extraction never saw this label.



Input set	Inputs	Macro F1	AUC	Fatal AUC
Majority baseline	none	0.24	0.50	0.50
Before the accident	88	0.48	0.70	0.74
Full chain	106	0.53	0.75	0.77
Full chain with report text	106+	0.54	0.76	0.80
Severity read from the text	text	0.82	n/a	n/a

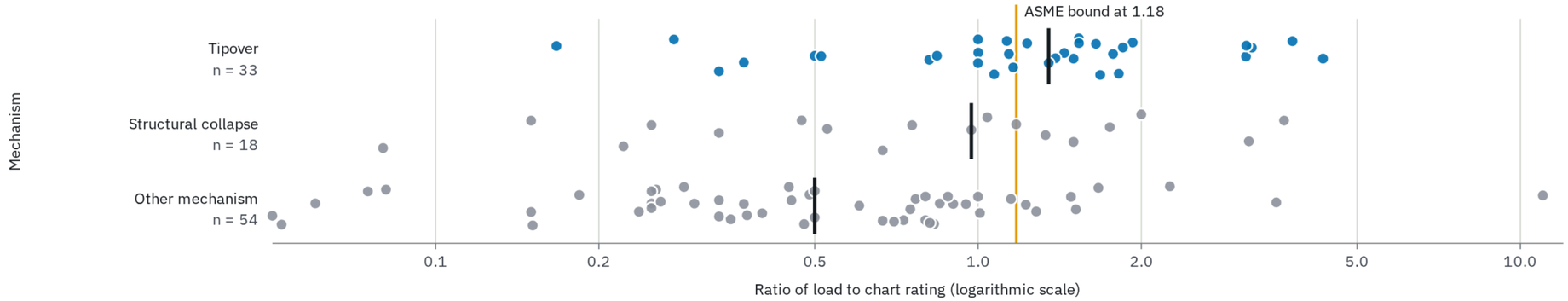
Best of eleven classifiers, 5-fold cross-validation repeated twice. The text model sees the narrative with every injury sentence removed.

> The chain predicts severity only in part. Where a person stood or which body part was hit rarely appears in the causal factors.

BACKUP · LOAD RATIO CHECK

Why do some cranes tip below 1.18 times their rating, even below 0.5?

Low-Ratio Tipovers Break the Chart's Assumptions



105 reports state both numbers. Tipover median 1.35, collapse 0.97, other 0.50, AUC 0.76. Of 33 tipovers, 15 lie below 1.18 and 4 below 0.5. None of the 15 is a plain static overload.

SETUP OR GROUND

5 ground failures or soft ground, 5 outriggers not set as charted, 1 slope

DYNAMIC OR SIDE LOAD

5 side pulls, travel with the load, or a sudden load release

RATING MISMATCH

1 limiter set for the wrong counterweight, about 4 nameplate values read as ratings

ASME B30.5, Mobile and Locomotive Cranes. One report can name several routes. Seven low ratios lie between 1.00 and 1.16.

DISCUSSION

Limitations and Next Steps

Limitations

- Static, rigid simulator, calibrated on two cranes: a Liebherr reference crane and the Terex T775 of case 2
- P(death | tipover) is closer to an upper bound
- Tipovers are about 12% of fatal crane accidents

Next Steps

- Structural collapse simulator (17% of fatal accidents), in development
- Load drop (22%) and power line clearance (17%) models
- Match weather records to wind accidents
- Exposure data to estimate accident rates
- Contact the company and obtain its 701 forensic accident cases (Wiethorn, 2018)
- Draft the paper manuscript