

The Jump Paradox: What 12 Million Powerlifting Attempts Can and Cannot Say About Attempt Selection

Zakhar Sukhanov

Independent researcher · zexa1993@outlook.com · [LinkedIn](#) · [GitHub](#)

Code, data pipeline and calculator: github.com/Dek0rta/third-attempt · dek0rta.github.io/third-attempt

Draft of 13 October 2026

Abstract. In powerlifting each lift gets three attempts, and the lifter picks every weight. Using every competition attempt recorded in the public OpenPowerlifting database from 2000 to 2026 (12,163,164 attempts by 572,553 lifters at 32,150 meets in 420 federations), I study how those choices relate to outcomes. Success falls from 86–95% on first attempts to 46–64% on third attempts. Openers above a lifter’s previous competition best are followed by a bomb-out (three misses) up to five times as often as openers at 85–92% of it. Success rates *rise* with the size of the jump up to about 6–7%. This “jump paradox” cannot be causal, since a heavier bar is never easier for the same lifter on the same day; it shows that lifters choose jumps using information about how they feel. A consequence is that a naive model fitted to these data recommends implausibly large jumps. Using bounds that assume only that heavier weights are never easier (monotone treatment response), I show that a third-attempt jump of 3–4.5% maximizes the expected gain that can be guaranteed, in all three lifts and across subgroups. Success models validated on 2025–26 meets are well calibrated but weakly discriminating (AUC 0.60–0.65, and at most 0.67 with the lifter’s own history). Misses earlier in a meet predict a shakier deadlift opener (odds $\times 0.76$ per miss on a first or second attempt) after adjusting for its weight and the day’s form. I release the code and a free calculator built on the models.

1 Introduction

A powerlifting meet gives each lifter three attempts in the squat, the bench press and the deadlift. The best successful attempt in each lift counts toward the total, and a lifter who misses all three attempts in any lift “bombs out” and has no total. Attempt selection is therefore a sequence of decisions under risk: a bigger jump adds more kilograms if it succeeds but is more likely to fail. Coaches rely on rules of thumb, such as opening at about 90% of the planned best and making smaller jumps on later attempts (Travis et al., 2021), but the sport now has a public record of millions of attempts that can test them.

Earlier work has described attempt success in samples of elite or single-federation lifters. Darragh et al. (2026) model success against jump size in kilograms for IPF raw lifters and report that moderate jumps are associated with higher success than small ones. Opener weight has been related to the chance of winning (Hoek et al., 2022; Howells et al., 2022). Travis et al. (2021) describe the attempt patterns of world-championship lifters who made all nine attempts. Nishihata & Otani (2024) show that competitive pressure lowers third-attempt bench success. In weightlifting, Genakos & Pagliero (2012) show that interim rank changes risk taking.

This paper makes four contributions. First, it uses all federations, equipment types and years that record attempts, and measures weights relative to each lifter’s previous competition best. Second, it documents a *jump paradox*, explains why it is evidence of selection rather than a dose–response curve, and shows how it misleads naive models. Third, it answers the decision question that matters to lifters, which third-attempt jump to take, with partial-identification bounds (Manski, 1997) that do not depend on how lifters choose. Fourth, it measures how predictable attempts are at all and tests whether misses carry over to later lifts, a question related to the “hot hand” debate (Bar-Eli et al., 2006; Gilovich et al., 1985; Miller & Sanjurjo, 2018).

2 Data

2.1 Source and processing

OpenPowerlifting is a volunteer-maintained, public-domain archive of meet results (*Openpowerlifting: A Public-Domain Archive of Powerlifting History. Data Snapshot of 10 October 2026*, 2026). I use the snapshot of 10 October 2026, which contains 4,046,314 lifter-meet rows. A recorded attempt weight is positive if the lift was made and negative if it was missed. Not all federations publish individual attempts, so I keep only lift entries where all three attempts are recorded.

The same performance often appears several times because a lifter competed in several divisions at one meet; I keep one copy of each (lifter, meet, lift, attempts) combination. I drop 142 meets with ten or more entries and not a single recorded miss, since those federations evidently report only successful attempts. The analysis sample keeps meets from 2000 onward, men and women, known equipment, weights between 15 and 600 kg, and jumps between 0% and 30%.

2.2 Variables

For each attempt I compute the jump over the previous attempt in percent and whether the previous attempt was made. The lifter's *previous best* is their best successful competition lift in the same lift and equipment group (raw, wraps or equipped) at any strictly earlier meet in the database; 64% of attempts have one. *Experience* is the number of distinct earlier meet dates. Weights relative to the previous best are the key normalization: they make a 100 kg and a 300 kg lifter comparable without using anything that happened on the day.

The final sample has 12,163,164 attempts in 4,068,082 lift entries by 572,553 lifters at 32,150 meets held in 125 countries.

Lift	1st attempt	2nd attempt	3rd attempt
Squat	86.3%	81%	63.9%
Bench press	88.4%	76.8%	45.6%
Deadlift	94.7%	87.6%	59.6%

Table 1: Share of attempts made, 2000–2026.

3 What happens on the platform

Success falls with every attempt (Table 1, Figure 1). Third attempts are made 64% of the time in the squat, 60% in the deadlift and only 46% in the bench press. At the same weight relative to the previous best, later attempts fail more often (Figure 2). The fall is steepest in the deadlift, the last lift of the day.

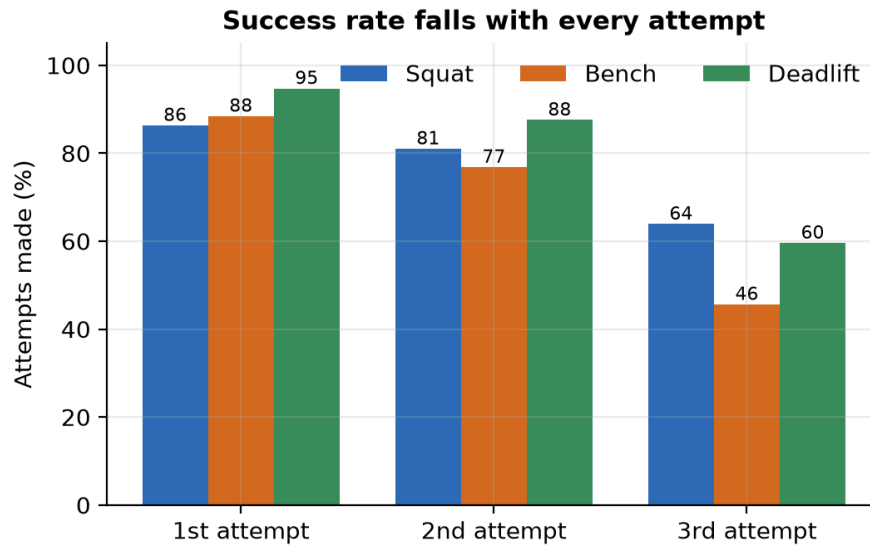


Figure 1: Share of attempts made by lift and attempt number.

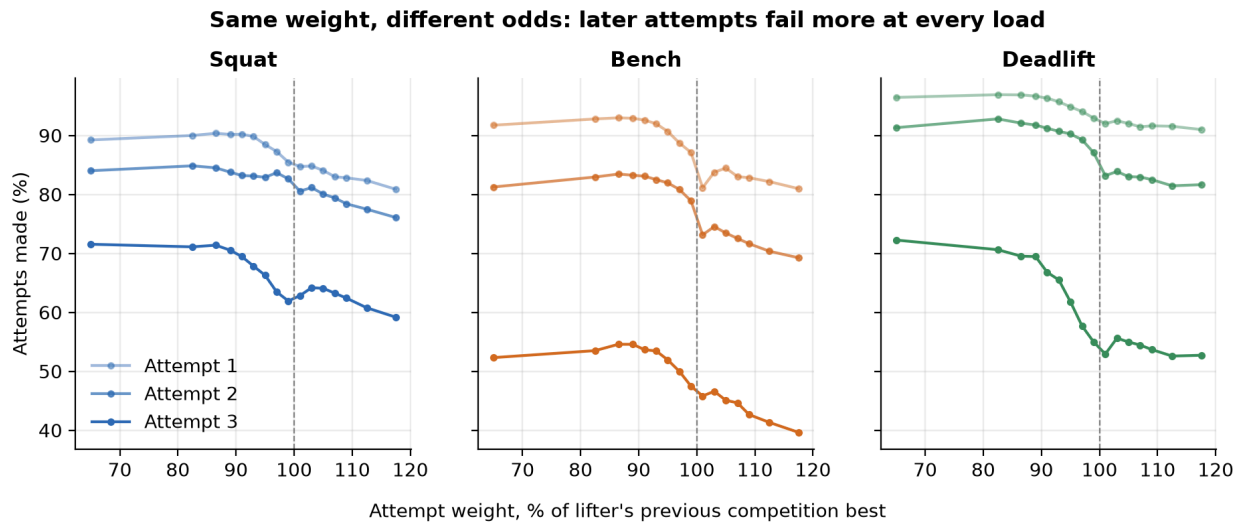


Figure 2: Share made by attempt weight relative to the lifter's previous competition best. Bins with at least 500 attempts.

Openers and bomb-outs. The median opener is 93–95% of the previous best. Bomb-out risk is flat at about 1% (squat and bench) and 0.3% (deadlift) for openers up to 92% of the previous best, then climbs steeply (Figure 3). Bench openers just above the previous best are followed by a bomb-out 5.8% of the time, against 1.0–1.1% at 85–92%. In the deadlift the ratio is similar (1.5–1.6% against 0.24–0.34%); in the squat it is smaller (2.6–3.2% against 1.0%).

This comparison is observational, but the direction of any bias is known. Lifters who open above their previous best are the ones who expect to have improved. If anything, that selection makes heavy openers look *safer* than they would be for a typical lifter, so the true increase in risk is likely larger than the one shown.

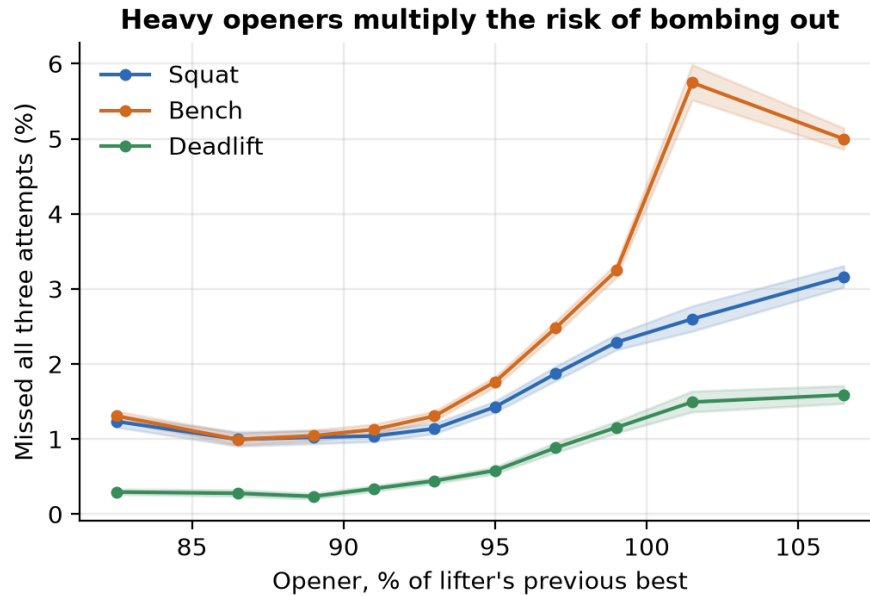


Figure 3: Share of lift entries with three missed attempts, by opener weight relative to the previous best. Shaded: 95% Wilson intervals.

4 The jump paradox

Figure 4 plots success against the jump over a made previous attempt. Between jumps of 0–2% and 6–7%, success on second attempts rises from 69% to 81% in the bench press, from 73% to 85% in the squat and from 73% to 89% in the deadlift. Third-attempt squats and deadlifts show the same rise.

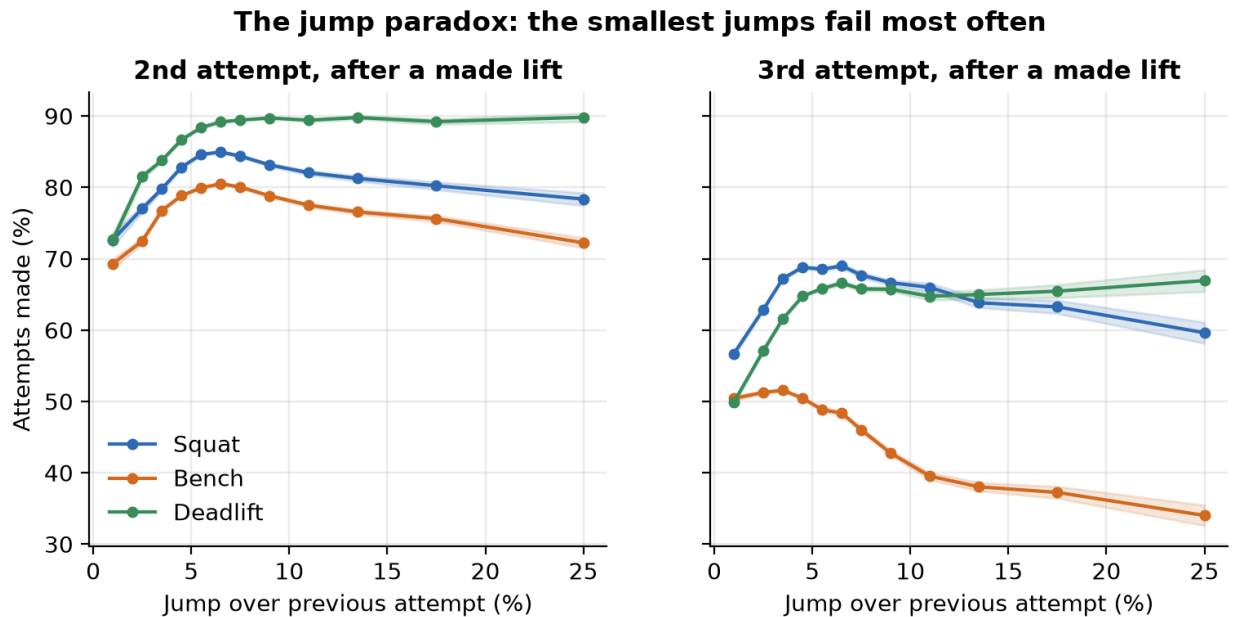


Figure 4: Share made by jump size after a made previous attempt. Shaded: 95% Wilson intervals; bins with at least 1,000 attempts.

This cannot be a causal effect. For one lifter on one day, adding weight to the bar can only make a lift harder. The rising part of the curve must come from *who* chooses each jump. A lifter whose previous attempt moved slowly takes a small jump or none; a lifter whose previous attempt flew takes a large one. The jump encodes private information about the day, so a success-by-jump curve is a mixture of different lifters, not a dose–

response curve. The same selection appears after a miss: going up after a missed attempt “succeeds” more often than repeating the weight (for example 45% against 34% on third-attempt deadlifts), most likely because lifters go up when the miss was technical rather than a lack of strength (Figure 5).

Any analysis that reads these curves as advice inherits the bias. To show how large it is, I fitted the success model of Section 6, plugged every jump from 0.5% to 15% into it for real 2025–26 lifters, and maximized expected gain. In all three lifts the “optimal” third-attempt jump was the largest one allowed, 15%, because predicted success falls more slowly than the jump grows. That answer is an artifact of selection.

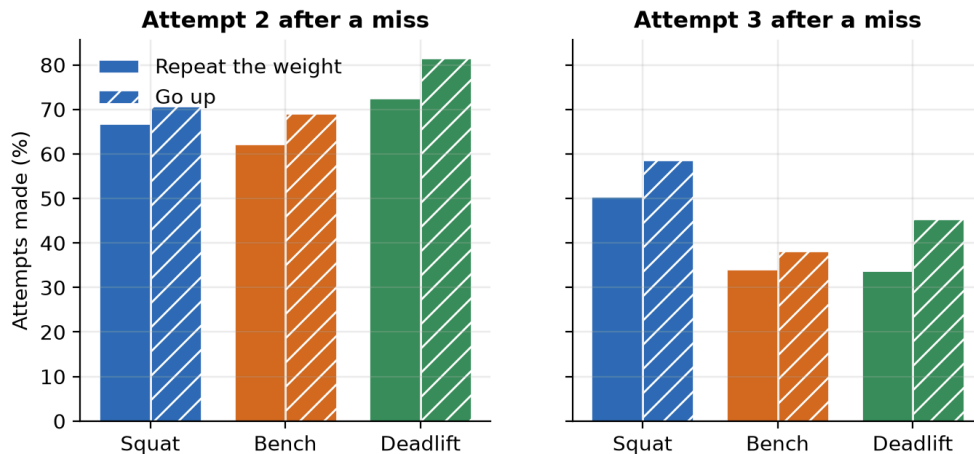


Figure 5: Share made after a missed attempt, by whether the lifter repeated the weight or went up.

5 Choosing the third attempt without assuming how lifters choose

5.1 Bounds under monotone treatment response

Instead of a model of how lifters choose jumps, I use a single assumption: for a given lifter on a given day, if a weight would be made, any lighter weight would also be made. This is monotone treatment response (MTR) (Manski, 1997). Consider third attempts after a made second attempt, group jumps into bins, and ask what share of *all* these lifters would have made jump t .

For a lifter who chose a bigger jump than t and made it, MTR says they would also have made t . For a lifter who chose a smaller jump and missed, MTR says they would also have missed t . Everyone else is unknown and could go either way. Writing p_u for the share choosing bin u and m_u for their success rate,

$$\sum_{u \geq t} p_u m_u \leq P(\text{make} \mid t) \leq 1 - \sum_{u \leq t} p_u (1 - m_u).$$

Multiplying by t gives bounds on the expected gain of jump t , in percent of the second attempt. These bounds do not depend on how lifters picked their jumps.

5.2 Result: a 3–4.5% third attempt has the best guaranteed gain

The bounds are wide, as they must be without further assumptions (Figure 6). The upper bound keeps rising with the jump, so the data alone cannot rule out that some lifters would gain from large jumps. The lower bound, however, has a clear peak. Following Manski’s treatment of decisions under ambiguity (Manski, 2000), I report the *maximin* jump: the one whose worst-case expected gain is highest. It falls in the 3.5–4.5% bin for the squat, the bench press and the deadlift, with a guaranteed gain of about 2.0%, 1.3% and 1.9% of the second attempt.

What the data can and cannot say about the third attempt

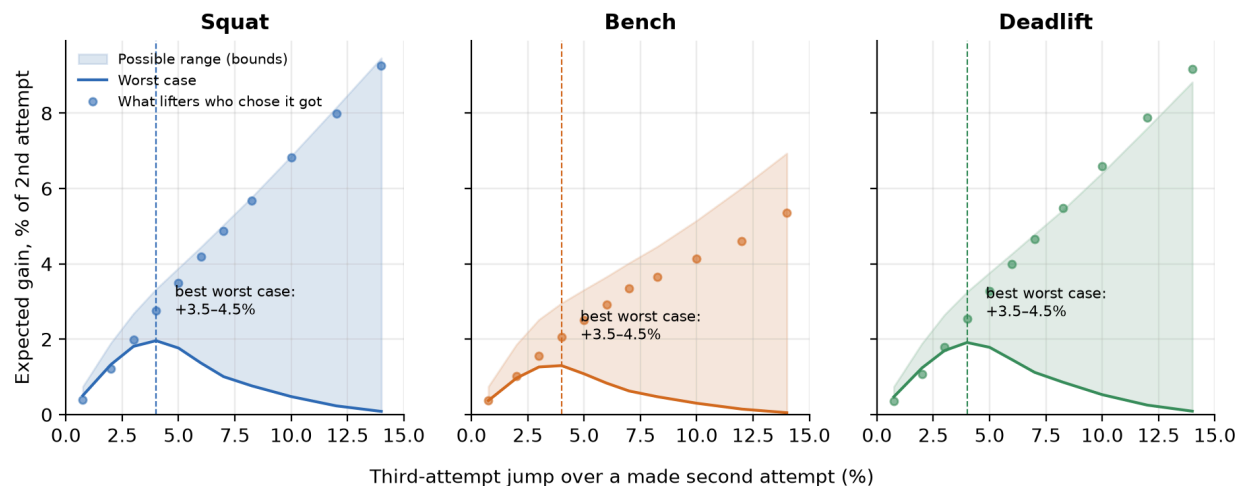


Figure 6: Expected gain of a third attempt by jump size. Shaded: range allowed by monotone treatment response. Line: worst case. Points: what lifters who chose each jump gained.

The result is robust (Table 2). On a finer 0.5% grid the maximin jump is 3.5–4% (squat), 3–3.5% (bench press) and 4–4.5% (deadlift). Within subgroups it stays between 2.5% and 4.5%; the only groups below 3.5% are male, equipped and very experienced benchers and very experienced squatters. A 3–4.5% third attempt is also close to what elite lifters do in practice (Travis et al., 2021).

Group	Squat	Bench	Deadlift
All (0.5% grid)	3.5–4%	3–3.5%	4–4.5%
Men	3.5–4.5%	2.5–3.5%	3.5–4.5%
Women	3.5–4.5%	3.5–4.5%	3.5–4.5%
Raw	3.5–4.5%	3.5–4.5%	3.5–4.5%
Wraps	3.5–4.5%	3.5–4.5%	3.5–4.5%
Equipped	3.5–4.5%	2.5–3.5%	3.5–4.5%
0-1 prior meets	3.5–4.5%	3.5–4.5%	3.5–4.5%
8+ prior meets	2.5–3.5%	2.5–3.5%	3.5–4.5%
2020 and later	3.5–4.5%	3.5–4.5%	3.5–4.5%
IPF affiliates	3.5–4.5%	3.5–4.5%	3.5–4.5%

Table 2: Maximin third-attempt jump bin by subgroup (percent over a made second attempt).

Maximin is a cautious rule: it protects against the worst case allowed by the data, not the most likely one. A lifter who knows their second attempt was easy has exactly the information that the bounds refuse to assume, and may rationally jump more. The contribution here is a defensible default and a clear statement of what the data alone can support.

6 How predictable is an attempt?

6.1 Models

I fit one logistic regression per lift and attempt number, separately for lifters with and without a previous best. Features are linear splines in the weight relative to the previous best (knots at 85–110%) and in the jump (knots at 2–15%), the previous result, the first-attempt result and first-to-second jump (for third attempts), sex, equipment, drug testing, single-lift meets, experience bins, age bins and the 40 largest federations. Linear splines keep the model exactly reproducible in a browser; a test in the repository checks that the JavaScript and Python predictions agree to 10^{-9} on 720 real situations.

Models are trained on meets before 2025 and evaluated on 2025–26 meets they never saw. As a ceiling, I also add the lifter’s own make rate over all earlier attempts, shrunk toward 75%.

6.2 Calibrated, but weakly discriminating

Model	Test attempts	AUC basic	+ federation	+ own history	Brier skill
Squat, attempt 1	149,044	0.593	0.599	0.624	0.018
Squat, attempt 2	149,044	0.621	0.624	0.645	0.036
Squat, attempt 3	149,044	0.598	0.599	0.616	0.031
Bench, attempt 1	182,959	0.638	0.643	0.661	0.043
Bench, attempt 2	182,959	0.616	0.620	0.641	0.046
Bench, attempt 3	182,959	0.598	0.603	0.618	0.037
Deadlift, attempt 1	158,850	0.623	0.634	0.649	0.012
Deadlift, attempt 2	158,850	0.645	0.649	0.665	0.034
Deadlift, attempt 3	158,850	0.641	0.646	0.657	0.066

Table 3: Out-of-time performance on 2025–26 meets, lifters with a previous best. Brier skill is relative to the training-set success rate; it is reported for the federation model used by the calculator.

The predicted probabilities are well calibrated on the unseen meets (Figure 7): when the model says 60%, close to 60% of attempts are made. But the models separate makes from misses only weakly. For lifters with a previous best, the AUC is 0.60–0.65 with the calculator’s inputs and at most 0.67 with the lifter’s own history (Table 3). Most of what decides an attempt is not in the public record: how the warm-ups went, technique on the day, fatigue, and judging.

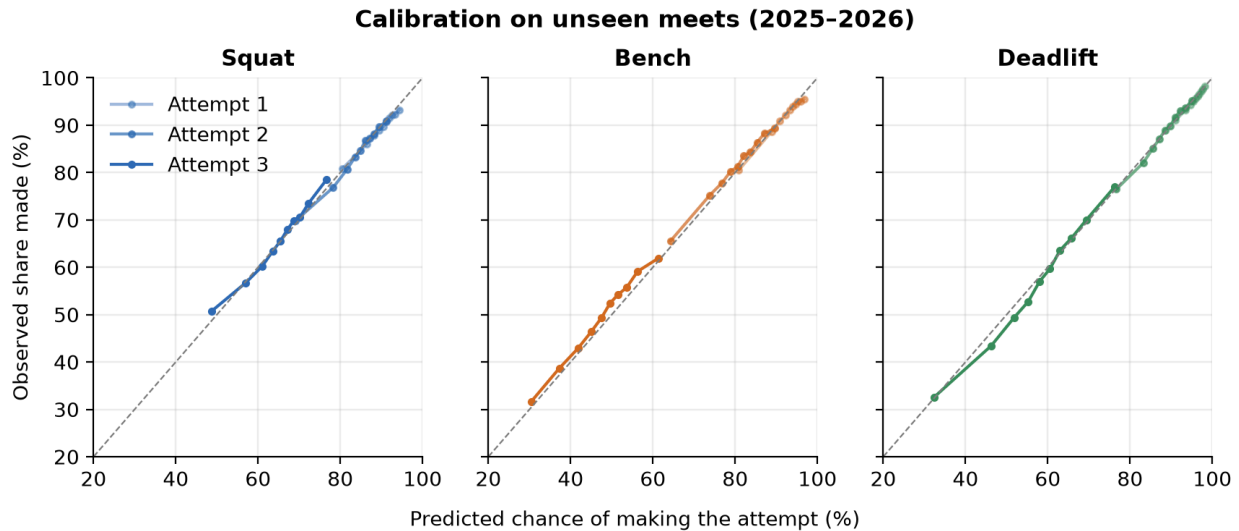


Figure 7: Calibration on 2025–26 meets: deciles of predicted probability against the observed share made.

7 Do misses carry over?

The hot-hand literature asks whether outcomes cluster beyond chance (Gilovich et al., 1985; Miller & Sanjurjo, 2018). In powerlifting the cleanest version of the question is across lifts: does missing squats and bench presses predict missing the deadlift opener, a weight usually planned before the meet and made 95% of the time?

I use full-power entries where the lifter has a previous best in all three lifts and made at least one squat and one bench press (642,072 entries). The outcome is the deadlift opener. Controls are its weight relative to the previous best (a linear spline), sex, equipment, testing and experience. A second model adds the day’s form: the best squat and bench press of the day and the heaviest attempted weights, all relative to the previous bests. Standard errors are clustered by lifter.

Odds ratio per miss (95% CI)	Weight controls	+ day's form
Miss on a 1st or 2nd squat/bench attempt	0.75 (0.74–0.76)	0.76 (0.75–0.77)
Miss on a 3rd squat/bench attempt	0.85 (0.83–0.86)	0.82 (0.80–0.84)

Table 4: Logistic regression of making the deadlift opener on earlier misses.

Each early miss lowers the odds of making the deadlift opener by about a quarter (odds ratio 0.76), and each third-attempt miss by about a fifth (0.82), with or without the day's form (Table 4, Figure 8). Early misses matter more, which fits an information story: missing an easy attempt says more about the day than missing an ambitious one. These results show that misses predict later misses; they cannot separate a psychological effect from unmeasured fatigue, injury or a bad day.

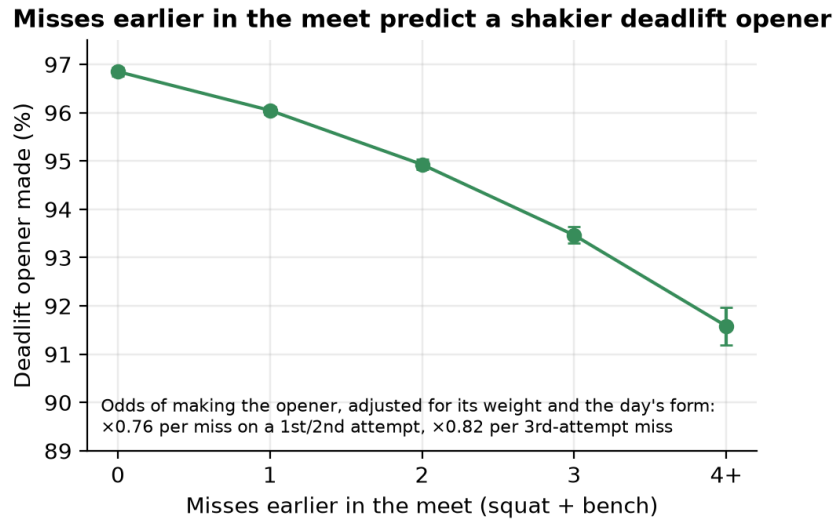


Figure 8: Share of deadlift openers made by the number of earlier misses, with 95% Wilson intervals.

8 A calculator that does not overclaim

The models power a free web calculator ([repository](#)). A lifter enters their previous best and their attempts so far; for every weight they could call next, it shows how often lifters in that situation made it, and for openers, the bomb-out risk when repeating after misses. That risk chains three model predictions; on 2025–26 meets its average (1.42%, 1.56%, 0.64% for squat, bench press and deadlift) is close to, and slightly above, the observed rate (1.25%, 1.43%, 0.56%). Two design choices follow from this paper. The calculator states every number as a description (“*N* of every 100 lifters in your spot made it”) rather than a prediction for anyone who forces a weight. And it does not show an “expected gain” column, which would have reproduced the selection artifact of Section 5; instead it marks how often lifters actually choose each jump, warns when a jump is rare, and highlights the 3–4.5% band of Section 5.

9 Limitations

The data are hand-entered from meet results and contain errors; the filters remove implausible values but not all mistakes. Federations that publish attempts may differ from those that do not. The previous best only covers meets in the database, so some lifters’ true bests are higher. Weights are grouped into three equipment classes. The MTR assumption treats each lifter-day as having a fixed capacity; random technical faults weaken it slightly. With millions of observations almost every difference is statistically significant, so the paper emphasizes sizes and directions over p-values. Finally, none of the results identify psychological mechanisms.

10 Conclusion

Public meet records can answer some attempt-selection questions and not others. They show clearly that later attempts fail more, that heavy openers raise bomb-out risk, and that misses travel through a meet. They cannot be read as dose–response curves for jump size, because lifters choose jumps with information the data do not contain; the jump paradox is the visible sign of that. With only the assumption that heavier is never easier, a third-attempt jump of 3–4.5% is the choice with the best guaranteed expected gain. All code, results and the calculator are open.

Acknowledgements. Thanks to the OpenPowerlifting contributors who keep a sport’s history public.

References

- Bar-Eli, M., Avugos, S., & Raab, M. (2006). Twenty years of “hot hand” research: Review and critique. *Psychology of Sport and Exercise*, 7(6), 525–553.
- Darragh, I. A. J., Egan, B., Nolan, D., & Bennett, K. E. (2026). Predicting a Successful Attempt in Raw Powerlifting. A Nonlinear Mixed Logistic Regression Analysis. *Journal of Strength and Conditioning Research*, 40(2), e138–e146. <https://doi.org/10.1519/JSC.0000000000005276>
- Genakos, C., & Pagliero, M. (2012). Interim Rank, Risk Taking, and Performance in Dynamic Tournaments. *Journal of Political Economy*, 120(4), 782–813.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314.
- Hoek, D. J. van den, Owen, P. J., Garrett, J. M., Howells, R. J., Pearson, J., Spathis, J. G., & Latella, C. (2022). What are the odds? Identifying factors related to competitive success in powerlifting. *BMC Sports Science, Medicine and Rehabilitation*, 14, 110. <https://doi.org/10.1186/s13102-022-00505-2>
- Howells, R. J., Spathis, J. G., Pearson, J., Latella, C., Garrett, J. M., Owen, P. J., & Hoek, D. J. van den. (2022). Impacts of squat attempt weight selection and success on powerlifting performance. *The Journal of Sports Medicine and Physical Fitness*, 62(4), 476–484. <https://doi.org/10.23736/S0022-4707.21.12140-1>
- Manski, C. F. (1997). Monotone Treatment Response. *Econometrica*, 65(6), 1311–1334.
- Manski, C. F. (2000). Identification problems and decisions under ambiguity: Empirical analysis of treatment response and normative analysis of treatment choice. *Journal of Econometrics*, 95(2), 415–442.
- Miller, J. B., & Sanjurjo, A. (2018). Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers. *Econometrica*, 86(6), 2019–2047.
- Nishihata, M., & Otani, S. (2024). *Reference Points, Risk-Taking Behavior, and Competitive Outcomes in Sequential Settings*. <https://doi.org/10.48550/arXiv.2409.13333>
- OpenPowerlifting: a public-domain archive of powerlifting history. Data snapshot of 10 October 2026.* (2026,). <https://www.openpowerlifting.org/>
- Travis, S. K., Zourdos, M. C., & Bazyler, C. D. (2021). Weight Selection Attempts of Elite Classic Powerlifters. *Perceptual and Motor Skills*, 128(1), 507–521. <https://doi.org/10.1177/0031512520967608>