Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Type 2 Diabetes, Metformin and Prediabetes: A Results Sheet of What Raises Risk, What Lowers Blood Sugar and What People Search

Narendra Kumar

b2b@fliarbi.com

Independent Researcher https://orcid.org/0009-0001-7468-6100

Research Article

Keywords: metformin, type 2 diabetes, diabetes medication, blood sugar levels, prediabetes, type 2 diabetes symptoms, type 2 diabetes and weight loss, type 2 diabetes treatment

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11753556/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare potential competing interests as follows: The author is associated with Fliar BI, whose data platform (Prism) supplied the records analysed in this study. No funding was received for this work, and there are no other competing interests.

RESEARCH PREPRINT · RESULTS SHEET · NOT PEER REVIEWED · NOVEMBER 2026

Type 2 Diabetes, Metformin and Prediabetes: A Results Sheet of What Raises Risk, What Lowers Blood Sugar and What People Search

A laboratory-style reading of 70,675 papers, 18,129 news articles and twenty-two years of search interest, flagged against a reference range

Narendra Kumar · ORCID 0009-0001-7468-6100 · Budaun, India · b2b@fliarbi.com

SPECIMEN research findings	RECORDS 125,816 findings / 70,675 papers	WINDOW 2017-2025 research; 2021-2026 news	PANELS A to H
REFERENCE RANGE ratio 1.00	FLAGS HIGH / LOW / WITHIN	SECOND CHECK precise half	VALIDATION blind reader, 70%

Abstract

A pooled reading of the research, news and search record on type 2 diabetes, laid out as a results sheet. Each result carries a ratio, an interval and a robustness flag. Eight results:

1. **Risk.** Adiposity 2.17 (95% CI 1.80 to 2.60), insulin resistance 2.10 (95% CI 1.51 to 2.93), blood pressure 1.69 (95% CI 1.37 to 2.08) and liver fat 2.13 (95% CI 1.37 to 3.34) carry the largest ratios for new type 2 diabetes and hold in the more precise half of studies (k = 77, 11, 20, 11; adiposity prediction interval 0.45 to 10.52).
2. **Imprecision.** Less precise studies report larger ratios: x1.07 per 0.1 standard error for risk factors and x1.20 for diabetes as exposure; the least precise third pools to 2.70 against 1.49 for the most precise.
3. **Nulls and failures.** Physical activity (0.91, interval up to 1.00), chemicals (1.07) and male sex (1.32) are gone in the precise half; diet quality (0.80 to 1.03) and alcohol include 1.00; prediabetes and cardiovascular events (1.33) is gone in the precise half.
4. **Consequences.** Type 2 diabetes raises later risk of kidney disease (3.14), cardiovascular events (1.67), death (1.47), dementia (1.42) and cancer (1.57); only kidney disease grows in the precise half. Prediabetes sits below diabetes on all four shared outcomes.
5. **Treatment.** SGLT2 inhibitors (0.56), GLP-1 agonists (0.65) and metformin (0.79) pair with fewer deaths, all from observational studies; metformin and dementia (1.08) and GLP-1 and cancer (0.97) are nulls. Reviews report an HbA1c fall of 0.28 percentage points (95% CI 0.17 to 0.39), 0.20 in the more precise reviews.
6. **Prevalence.** Pooled adult diabetes 10.3% (95% CI 9.7 to 10.9; prediction interval 5.4 to 18.9%) and prediabetes 20.0%; studies differ more than regions do.
7. **Context: attention.** Newer drugs are 5.7 times older drugs in validated news flags and 4.9 times in search, against 0.51 in research; news tracks search month to month only for Ozempic (r = 0.38).
8. **Context: counts and growth.** 35,903 typed papers name diabetes or prediabetes; the news rate is 1.68 times its 2018-2019 level; the anchor search term is 1.91 times its 2019 level against 1.13 for the control.

Keywords: metformin; type 2 diabetes; diabetes medication; blood sugar levels; prediabetes; type 2 diabetes symptoms; type 2 diabetes and weight loss; type 2 diabetes treatment

In plain words

In a sentence: Excess weight, insulin resistance, high blood pressure and fatty liver are the strongest links to getting type 2 diabetes, but the headline sizes are inflated by small studies, and the drug results come from observational data, not trials.

- **Weight is the big one.** Excess body fat roughly doubles the odds of new type 2 diabetes (about 2.2 times; 1.9 times in the more careful studies). Insulin resistance (2.1), fatty liver (2.1) and high blood pressure (1.7) are in the same band.
- **Small studies exaggerate.** For what diabetes leads to, the smallest third of studies pools to 2.7 times and the most careful third to 1.5 times. Trust the smaller number.

- **Some familiar claims did not hold up.** Physical activity (0.91), air pollution and chemicals (1.07) and being male (1.3) lose their effect once the careful studies are weighed. Alcohol is a null with a wide range. Better diet quality looked stronger, not weaker, in the careful studies (0.82).
- **Diabetes raises later risk, most for the kidneys.** Kidney disease is about 3.1 times, heart events 1.7 (1.5 in careful studies), death 1.5, cancer 1.6 and dementia 1.4. Prediabetes sits lower on all four shared outcomes, and its heart-event link is not clear in careful studies.
- **Drugs go with fewer deaths, but linked with is not the same as caused by.** SGLT2 drugs 0.56, GLP-1 drugs 0.65 and metformin 0.79 times the odds of death, all from studies that watch people. Metformin and dementia (1.08) and GLP-1 drugs and cancer (0.97) show no clear effect.
- **Reviews report a modest blood-sugar fall.** HbA1c drops by about 0.28 points on average, and about 0.20 in the more careful reviews, less than the 0.5 points often seen as worthwhile.
- **Attention is not evidence.** Newer drugs are about 5.7 times more common in news than metformin and 4.9 times in search, while this research set has about half as many GLP-1 studies as metformin studies. About 10% of adults in the studied surveys have diabetes; searches for the topic have nearly doubled since 2019, with a 2026 surge nobody can yet explain.

How to read the numbers

- **A ratio of 1.00 means no difference.** 1.50 means 50% higher odds, 0.80 means 20% lower.
- **Precise means big and careful.** Small studies are noisy, so their results swing wide, and the ones that get noticed are often the dramatic ones.
- **The 95% range (CI) is where the true figure most likely sits.** If it includes 1.00, the study cannot say there is any effect.
- **k is the number of studies pooled.** The fewer the studies, the shakier the pooled figure.
- **Null means we looked and found nothing clear.** It does not prove there is nothing there.
- **Linked with is not the same as caused by.** Most of these studies watch people; they do not run experiments.

1 Scope, specimen and how to read the sheet

HEADER

Question: which exposures raise the risk of type 2 diabetes, what does diabetes or prediabetes lead to, and what do glucose-lowering treatments change, once the studies are weighed by their precision? The sheet then sets the answer beside what news and search put in front of the public. Specimen: 125,816 extracted findings from 70,675 papers (2017 to 2025, each cited at least ten times), 18,129 news articles (January 2021 to April 2026, with a 2018 to 2020 baseline of 4,028) and monthly Google Trends interest from 2004 to 2026. Every ratio was re-read from the quoted sentence, and every field used in a claim was checked by a blind second reader (Section 11).

Reading a row. Result is a pooled ratio on the log scale (DerSimonian-Laird random effects (DerSimonian and Laird, 1986)). Reference range is a ratio of 1.00, meaning no association. Flag: `HIGH` when the 95% confidence interval lies above 1.00, `LOW` when it lies below, `WITHIN` when it includes 1.00. A second flag is the precise-half check: the studies are split at the median standard error and the pooled ratio is recomputed for the more precise half. `HOLDS` means that half still excludes 1.00 and lies within 1.5 times the imprecise half on the log scale; `SMALLER` means it holds but shrinks; `GONE` means its interval includes 1.00. The 95% prediction interval, from a t distribution with k minus 2 degrees of freedom, shows where a new study would fall (IntHout et al., 2016; Higgins et al., 2008).

Four lenses, never pooled together. RISK: an exposure and incident type 2 diabetes in people free of diabetes. OUT: type 2 diabetes as the exposure and a later outcome. PRE: prediabetes as the exposure. TRX: a glucose-lowering class and an outcome. Mean differences reported by reviews (HbA1c, HOMA-IR) and prevalence are separate analyses with their own units.

2 Findings

SUMMARY SHEET

Eight results. Results 1 to 6 are findings (effect, interval, robustness). Results 7 and 8 are context and sit last, as the series rule requires. Table 1 lists the headline rows; Figure 1 plots them.

F1 Adiposity, insulin resistance, blood pressure and liver fat carry the largest ratios for new type 2 diabetes, and they survive the precise-half check. Adiposity 2.17 (95% CI 1.80 to 2.60), k = 77, prediction interval 0.45 to 10.52; insulin resistance indices 2.10 (95% CI 1.51 to 2.93), k = 11; blood pressure 1.69 (95% CI 1.37 to 2.08), k = 20; liver fat 2.13 (95% CI 1.37 to 3.34), k = 11. For adiposity the more precise half gives 1.92. Flag: HIGH, HOLDS. Observational studies throughout, so these are associations.

F2 The less precise the study, the larger the ratio. Across 485 risk ratios the ratio multiplies by 1.07 (95% CI 1.02 to 1.13) for each 0.1 rise in standard error, within family. For type 2 diabetes as exposure the multiplier is 1.20 (1.16 to 1.24). The most precise third of rows pools to 1.32 against 1.57 for the least precise third; for the diabetes-as-exposure lens the thirds give 1.49 and 2.70. Treat pooled headline ratios as upper bounds.

F3 Three common claims do not hold up, one is a null, and diet quality gets stronger with precision. Physical activity 0.91 (95% CI 0.82 to 1.00) is gone in the precise half (0.95, interval 0.87 to 1.04). Chemicals and air pollution (1.07; precise half 1.03) and male sex (1.32; precise half 1.14) are gone in the precise half. Alcohol intake (1.26 (95% CI 0.79 to 2.00), k = 7) is a null with a wide interval. Diet quality 0.90 (95% CI 0.80 to 1.03) includes 1.00 overall, yet its more precise half is 0.82 (95% CI 0.78 to 0.86) and holds.

F4 Type 2 diabetes raises later risk, most for the kidney. Kidney disease 3.14 (95% CI 2.32 to 4.26), cardiovascular events 1.67 (95% CI 1.55 to 1.78), death 1.47 (95% CI 1.37 to 1.58), dementia 1.42 (95% CI 1.29 to 1.55), cancer 1.57 (95% CI 1.43 to 1.72). All stay HIGH in the precise half, but the ratios shrink for cardiovascular events (1.49 in the more precise half and 2.18 in the less precise half), death (1.31 in the precise half), cancer (1.31) and dementia (1.33); only kidney disease grows (3.46). Prediabetes as exposure: death 1.26 (95% CI 1.07 to 1.48) holds; cardiovascular events 1.33 (95% CI 1.13 to 1.56) is gone in the precise half.

F5 Glucose-lowering classes pair with fewer deaths, but the ratios come from observational data. SGLT2 inhibitors and death 0.56 (95% CI 0.44 to 0.72), k = 15; GLP-1 agonists and death 0.65 (95% CI 0.60 to 0.69), k = 5; metformin and death 0.79 (95% CI 0.70 to 0.88), k = 14. 139 of 142 class-to-outcome ratios come from observational studies and none from a trial. Nulls: metformin and dementia 1.08 (95% CI 0.84 to 1.39); GLP-1 agonists and cancer 0.97 (95% CI 0.54 to 1.73).

F6 Reviews report a modest HbA1c fall, and the more careful reviews report less. Pooled over 49 review estimates, the difference is -0.28 (95% CI -0.39 to -0.17) percentage points, prediction interval -1.01 to 0.45; the more precise half gives -0.20. By group: lifestyle -0.27, supplements -0.48, drugs -0.24. These are reviews, not head-to-head trials, so the groups are not ranked.

F7 Context: attention runs ahead of the evidence on newer drugs. GLP-1 and SGLT2 medicines are 20.0% of validated news flags against 3.5% for metformin and older drugs (ratio 5.7), while the research layer holds 21 GLP-1 studies against 41 for metformin. Search for Ozempic and Mounjaro, set against the anchor in the same request, is 4.9 times metformin in 2024 and 2025, against 0.12 in 2019.

F8 Context: counts and growth. The typed research layer holds 35,903 papers that name diabetes or prediabetes, rising from 48.7 to 64.5 per 10,000 papers between 2017 and 2024. News volume per 100,000 articles is 1.68 times its 2018-2019 level. Search for the anchor term averages 1.91 times its 2019 level in 2025 and 2026, against 1.13 for the weather control.

Table 1. Headline rows. Ratios pooled on the log scale; k = number of study ratios; flag = 95% CI against 1.00; second flag = precise-half check (Section 1). Full panels in Tables 2 to 4.

Panel	Result	Ratio (95% CI)	95% prediction interval	k	Flag	Precise half
A risk	Adiposity	2.17 (1.80 to 2.60)	0.45 to 10.52	77	HIGH	HOLDS
A risk	Insulin resistance indices	2.10 (1.51 to 2.93)	0.62 to 7.18	11	HIGH	HOLDS
A risk	Blood pressure	1.69 (1.37 to 2.08)	0.65 to 4.37	20	HIGH	HOLDS
A risk	Protective foods	0.89 (0.83 to 0.95)	0.62 to 1.26	37	LOW	SMALLER
A risk	Physical activity (higher)	0.91 (0.82 to 1.00)	0.57 to 1.45	28	WITHIN	GONE
A risk	Diet quality (higher)	0.90 (0.80 to 1.03)	0.51 to 1.59	22	WITHIN	HOLDS
C diabetes as exposure	Kidney disease	3.14 (2.32 to 4.26)	0.76 to 13.06	21	HIGH	HOLDS
C diabetes as exposure	Cardiovascular events	1.67 (1.55 to 1.78)	1.01 to 2.73	79	HIGH	SMALLER
C diabetes as exposure	Death	1.47 (1.37 to 1.58)	0.78 to 2.76	100	HIGH	SMALLER
C prediabetes	Cardiovascular events	1.33 (1.13 to 1.56)	0.74 to 2.40	14	HIGH	GONE
C prediabetes	Death	1.26 (1.07 to 1.48)	0.76 to 2.09	9	HIGH	HOLDS
D class to outcome	SGLT2 to death	0.56 (0.44 to 0.72)	0.21 to 1.53	15	LOW	HOLDS
D class to outcome	GLP-1 to death	0.65 (0.60 to 0.69)	0.56 to 0.75	5	LOW	NOT SPLIT
D class to outcome	Metformin to death	0.79 (0.70 to 0.88)	0.51 to 1.21	14	LOW	SMALLER
D class to outcome	Metformin to dementia	1.08 (0.84 to 1.39)	0.47 to 2.51	6	WITHIN	NOT SPLIT

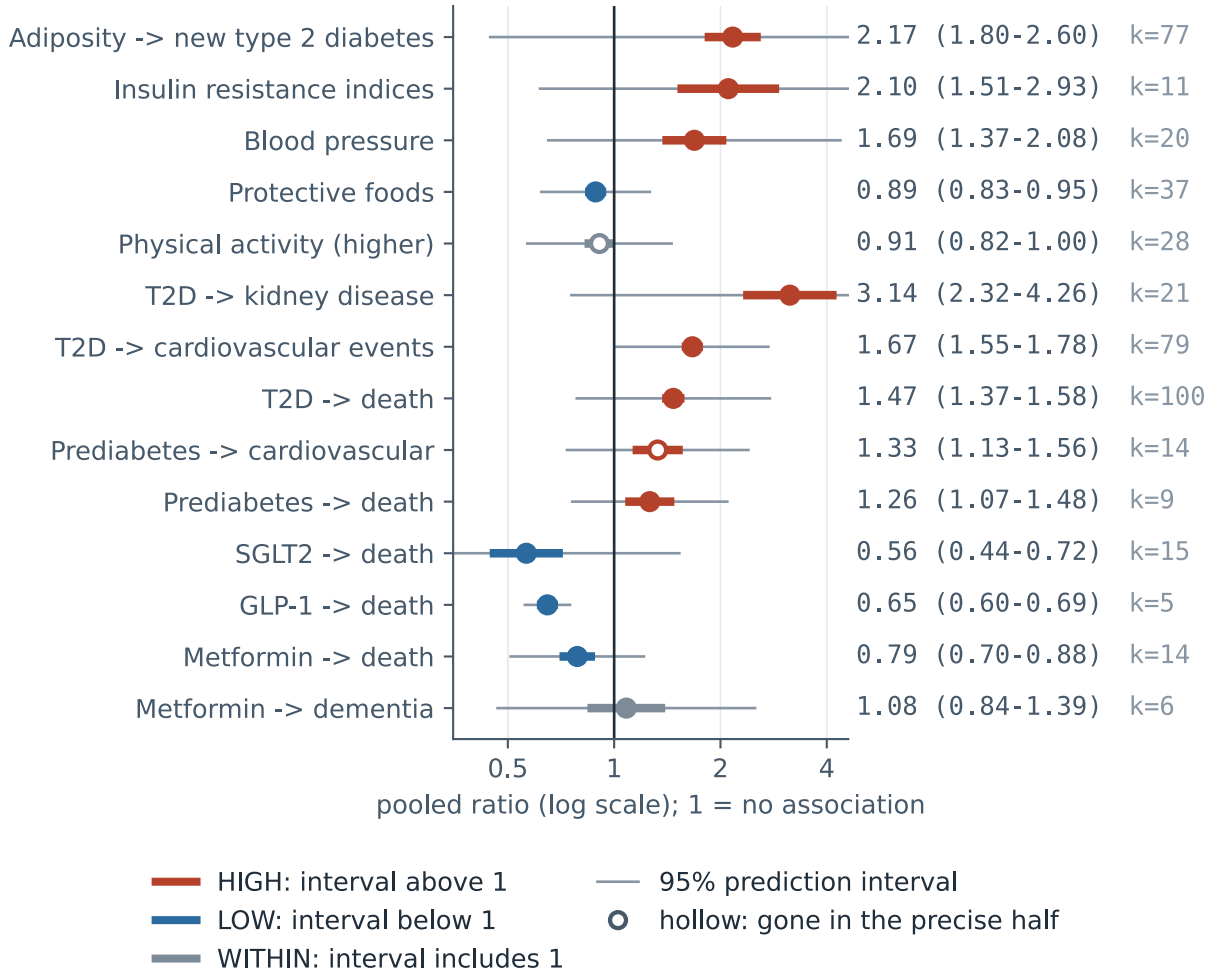


Figure 1. Summary sheet. Pooled ratio (bar: 95% CI; thin line: 95% prediction interval). Colour is the flag; a hollow marker means the ratio is gone in the precise half. Ratios for glucose-lowering classes are observational.

> INTERPRETATION Of the 15 headline rows, 12 are outside range. Of those, 6 hold in the precise half, 4 hold but shrink, 1 gone and 1 not split. The within-range rows are physical activity, diet quality and metformin with dementia. Every treatment ratio is observational.

3 Risk factors for new type 2 diabetes

PANEL A

Specimen: 485 ratios from 436 studies, one ratio per study and family, in people free of diabetes at baseline. 53% adjust for confounders; 257 are odds ratios and 193 hazard ratios. Overall pooled ratio 1.40 (95% CI 1.36 to 1.43), I-squared 99.8%. The overall figure mixes unlike exposures and is not a finding. Table 2 gives the panel by family; exposures are never pooled across families.

Table 2. Panel A. Risk factors for new type 2 diabetes. Precise / imprecise = pooled ratio in the more precise and the less precise half of studies.

Exposure family	k	Ratio (95% CI)	95% PI	I²	Precise / imprecise	Flag	Check
Adiposity (BMI, waist, fat)	77	2.17 (1.80 to 2.60)	0.45 to 10.52	100%	1.92 / 2.48	HIGH	HOLDS
Liver fat and enzymes	11	2.13 (1.37 to 3.34)	0.38 to 12.03	98%	2.33 / 1.78	HIGH	HOLDS
Insulin resistance (HOMA-IR, TyG)	11	2.10 (1.51 to 2.93)	0.62 to 7.18	98%	1.86 / 2.41	HIGH	HOLDS
Blood lipids	23	1.79 (1.37 to 2.34)	0.48 to 6.63	97%	1.49 / 2.21	HIGH	SMALLER
Blood pressure	20	1.69 (1.37 to 2.08)	0.65 to 4.37	98%	1.58 / 1.84	HIGH	HOLDS
Early-life and socioeconomic	44	1.65 (1.26 to 2.14)	0.28 to 9.72	100%	1.55 / 1.73	HIGH	HOLDS
Harmful diet (sugar drinks, processed meat)	11	1.62 (1.24 to 2.11)	0.59 to 4.44	99%	1.64 / 1.48	HIGH	HOLDS
Metabolic syndrome, uric acid	20	1.56 (1.27 to 1.92)	0.62 to 3.91	98%	1.37 / 1.84	HIGH	SMALLER
Short or poor sleep, apnoea	28	1.43 (1.24 to 1.65)	0.67 to 3.05	97%	1.34 / 1.58	HIGH	SMALLER
Male sex	12	1.32 (1.15 to 1.52)	0.81 to 2.17	96%	1.14 / 1.64	HIGH	GONE
Depression, gum disease, infections	34	1.32 (1.21 to 1.44)	0.86 to 2.03	96%	1.21 / 1.70	HIGH	SMALLER
Hormonal (PCOS, thyroid)	14	1.31 (1.18 to 1.46)	0.90 to 1.90	81%	1.30 / 1.33	HIGH	HOLDS
Sedentary time	8	1.29 (1.08 to 1.55)	0.75 to 2.22	81%	1.30 / 1.16	HIGH	HOLDS
Medicines (statins, steroids and others)	17	1.29 (1.19 to 1.39)	0.94 to 1.75	93%	1.24 / 1.33	HIGH	HOLDS
Alcohol intake	7	1.26 (0.79 to 2.00)	0.27 to 5.85	91%	-	WITHIN	NOT SPLIT
Smoking	8	1.25 (1.15 to 1.37)	0.95 to 1.65	85%	1.18 / 1.38	HIGH	SMALLER
Genetic variants, family history	32	1.24 (1.15 to 1.33)	0.94 to 1.63	92%	1.17 / 1.63	HIGH	SMALLER
Chemicals, metals, air pollution	21	1.07 (1.01 to 1.14)	0.86 to 1.33	85%	1.03 / 1.38	HIGH	GONE
Physical activity (higher)	28	0.91 (0.82 to 1.00)	0.57 to 1.45	95%	0.95 / 0.78	WITHIN	GONE
Diet quality (higher)	22	0.90 (0.80 to 1.03)	0.51 to 1.59	94%	0.82 / 0.94	WITHIN	HOLDS
Protective foods (coffee, nuts, fibre)	37	0.89 (0.83 to 0.95)	0.62 to 1.26	90%	0.92 / 0.77	LOW	SMALLER

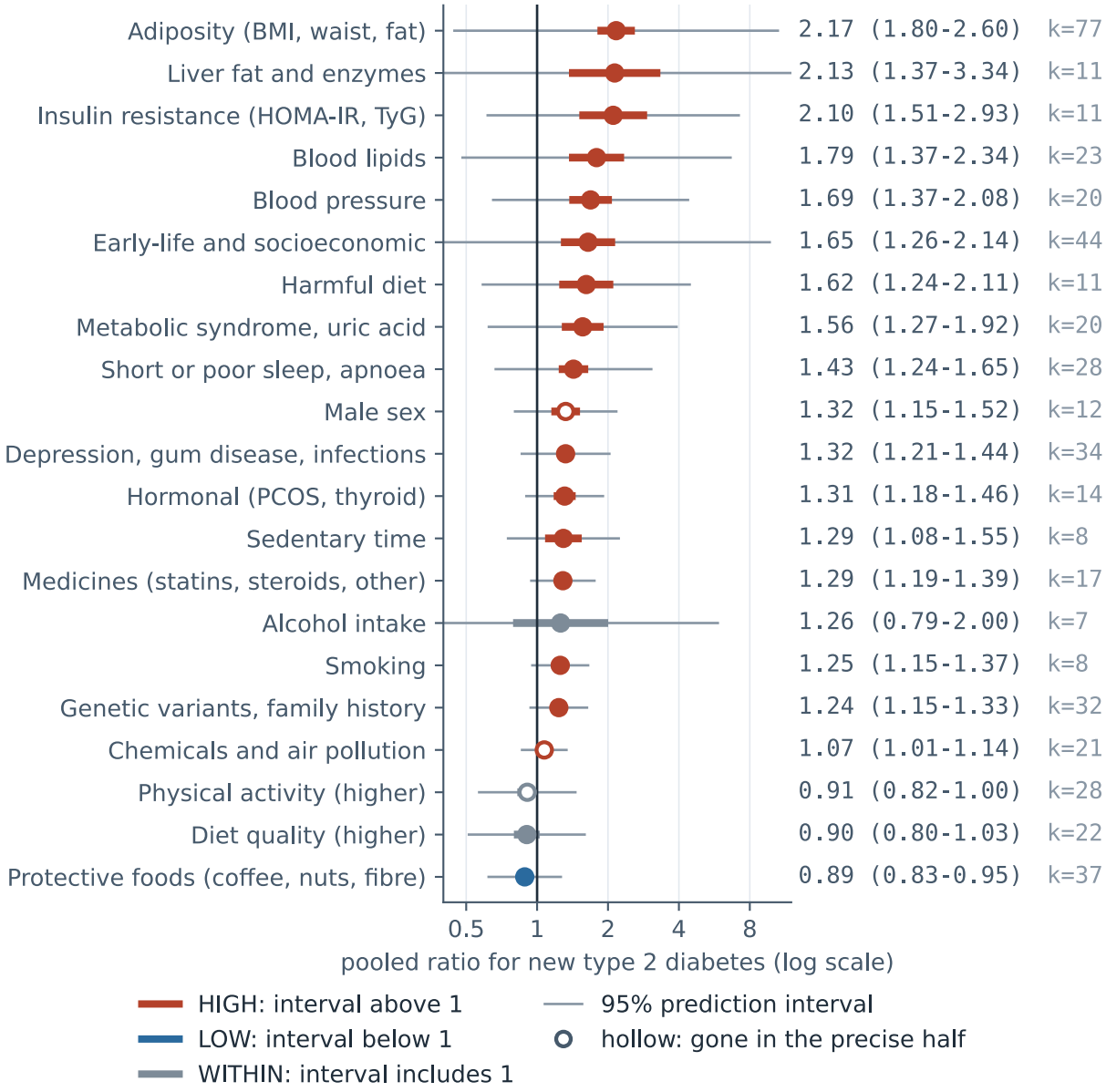


Figure 2. Panel A. Risk factors for new type 2 diabetes, sorted by ratio. Bar: 95% CI; thin line: 95% prediction interval; hollow marker: gone in the precise half.

Of 21 families, 17 hold in the precise half, 7 of them smaller; 3 are gone and 1 (alcohol) have too few ratios to split. Protective foods (0.89 (0.83-0.95)) are LOW but the precise half is 0.92 against 0.77 in the imprecise half. The sugar-sweetened drink and sedentary-time literatures behind the harmful-diet and sedentary rows are older (Malik et al., 2010; Aune et al., 2015), and the sheet agrees on direction.

Moderators were tested within family, with family dummies and the imprecision term. Adjusted against unadjusted: x0.98 (95% CI 0.88 to 1.10); 2021 or later: x0.96; odds ratio against hazard or risk ratio: x0.98. East Asian studies give x0.89 (p = 0.042), which does not survive a Bonferroni correction for 20 moderator tests (threshold 0.0025). No moderator is claimed.

> INTERPRETATION Adiposity and insulin resistance are outside range and robust. Physical activity is within range at the strict reading (1.001 is the upper limit) and gone in the precise half. The diet-quality row is within range overall, but the more precise half is below 1.00 and holds, so the overall figure may understate it. Heterogeneity is extreme everywhere (I-squared 81% to 100%), so single-family ratios describe a typical study, not a new population.

4 Precision check: how much of each ratio is imprecision

PANEL B

The precise-half check separates two readings of a large pooled ratio: a real effect, or small studies reporting dramatic numbers that get noticed (Ioannidis, 2008). Figure 3 sets the two halves side by side for every family with at least eight ratios. In the risk lens the less precise half is further from 1.00 in 16 of 20 split families; the others are sedentary time, diet quality, harmful diet, liver fat and enzymes. In the diabetes-as-exposure lens it is further from 1.00 in 9 of 10; the exception is kidney disease.

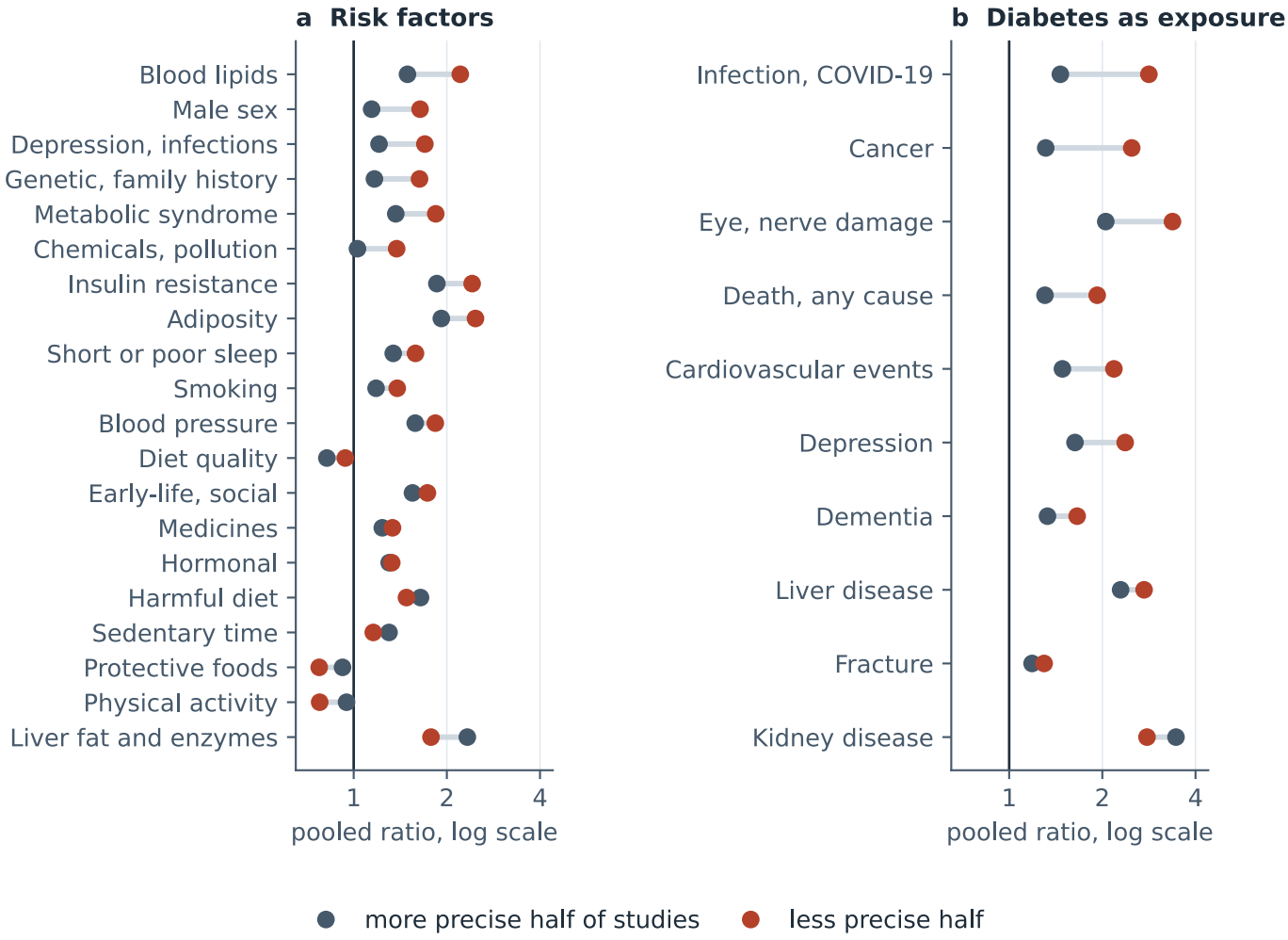


Figure 3. Panel B. Pooled ratio in the more precise half (dark) and less precise half (red) of studies, by family. Risk factors (a) and diabetes as exposure (b). A gap to the right means the less precise studies report larger ratios.

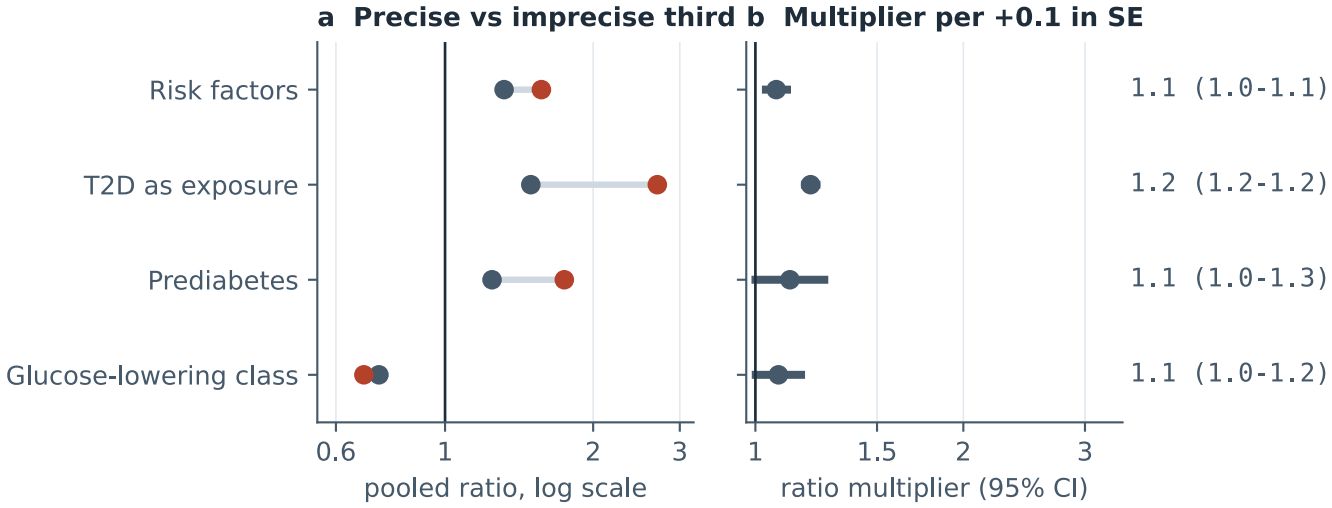


Figure 4. Panel B. (a) Pooled ratio in the most precise and least precise third of rows, by lens. (b) Multiplier on the ratio for each 0.1 rise in the standard error of the log ratio, within family (95% CI).

The multiplier survives alternatives. Cluster-robust by study, the RISK multiplier per unit standard error is 2.12 ($p = 0.014$) and for diabetes as exposure 5.09 ($p < 0.001$); varying the between-study variance from 0.02 to 0.3 gives 2.10 to 2.15 for RISK; leaving out each family in turn gives 1.93 to 2.46. For prediabetes ($k = 29$ ratios in 3 families) the multiplier is 1.12 with a 95% CI of 0.99 to 1.27, so the sign is not established; for glucose-lowering classes it is 1.08 (0.99 to 1.18), also not established. Funnel-style asymmetry in the pooled lenses is a pooled pattern, not a test of any one family (Egger et al., 1997; Sterne et al., 2011).

> INTERPRETATION Outside range, and partly an artefact. For diabetes as exposure the least precise third pools to 2.70 against 1.49; a reader should use the precise-half figure first.

5 What type 2 diabetes and prediabetes lead to

PANEL C

Specimen: 384 ratios from 353 studies for type 2 diabetes as exposure, and 40 ratios from 34 studies for prediabetes. Outcomes sit in listed families; all but six diabetes-as-exposure ratios are from observational designs. Comparisons are diabetes against none, not diabetes against better-controlled diabetes.

Table 3. Panel C. Type 2 diabetes (T2D) and prediabetes as exposures, by outcome family.

Outcome family	Exposure	k	Ratio (95% CI)	95% PI	Precise / imprecise	Flag	Check
Kidney disease	T2D	21	3.14 (2.32 to 4.26)	0.76 to 13.06	3.46 / 2.79	HIGH	HOLDS
Liver disease	T2D	21	2.55 (1.88 to 3.47)	0.61 to 10.74	2.29 / 2.73	HIGH	HOLDS
Eye and nerve damage	T2D	18	2.40 (1.89 to 3.04)	0.92 to 6.25	2.05 / 3.37	HIGH	SMALLER
Depression	T2D	9	1.92 (1.39 to 2.66)	0.64 to 5.79	1.63 / 2.37	HIGH	SMALLER
Severe infection, COVID-19	T2D	39	1.82 (1.60 to 2.07)	0.88 to 3.73	1.46 / 2.83	HIGH	SMALLER
Cardiovascular events	T2D	79	1.67 (1.55 to 1.78)	1.01 to 2.73	1.49 / 2.18	HIGH	SMALLER
Cancer (any site)	T2D	54	1.57 (1.43 to 1.72)	0.88 to 2.78	1.31 / 2.49	HIGH	SMALLER
Heart failure	T2D	6	1.49 (1.28 to 1.74)	0.96 to 2.33	-	HIGH	NOT SPLIT
Death, any cause	T2D	100	1.47 (1.37 to 1.58)	0.78 to 2.76	1.31 / 1.92	HIGH	SMALLER
Dementia and cognitive decline	T2D	22	1.42 (1.29 to 1.55)	0.97 to 2.07	1.33 / 1.66	HIGH	SMALLER
Fracture	T2D	9	1.20 (1.16 to 1.24)	1.13 to 1.27	1.19 / 1.30	HIGH	SMALLER
Atrial fibrillation	T2D	6	1.11 (1.02 to 1.21)	0.88 to 1.40	-	HIGH	NOT SPLIT
Kidney disease	Prediabetes	6	1.53 (1.08 to 2.16)	0.47 to 5.02	-	HIGH	NOT SPLIT
Cancer (any site)	Prediabetes	4	1.37 (1.10 to 1.72)	0.55 to 3.45	-	HIGH	NOT SPLIT
Cardiovascular events	Prediabetes	14	1.33 (1.13 to 1.56)	0.74 to 2.40	1.22 / 1.47	HIGH	GONE
Death, any cause	Prediabetes	9	1.26 (1.07 to 1.48)	0.76 to 2.09	1.32 / 1.24	HIGH	HOLDS

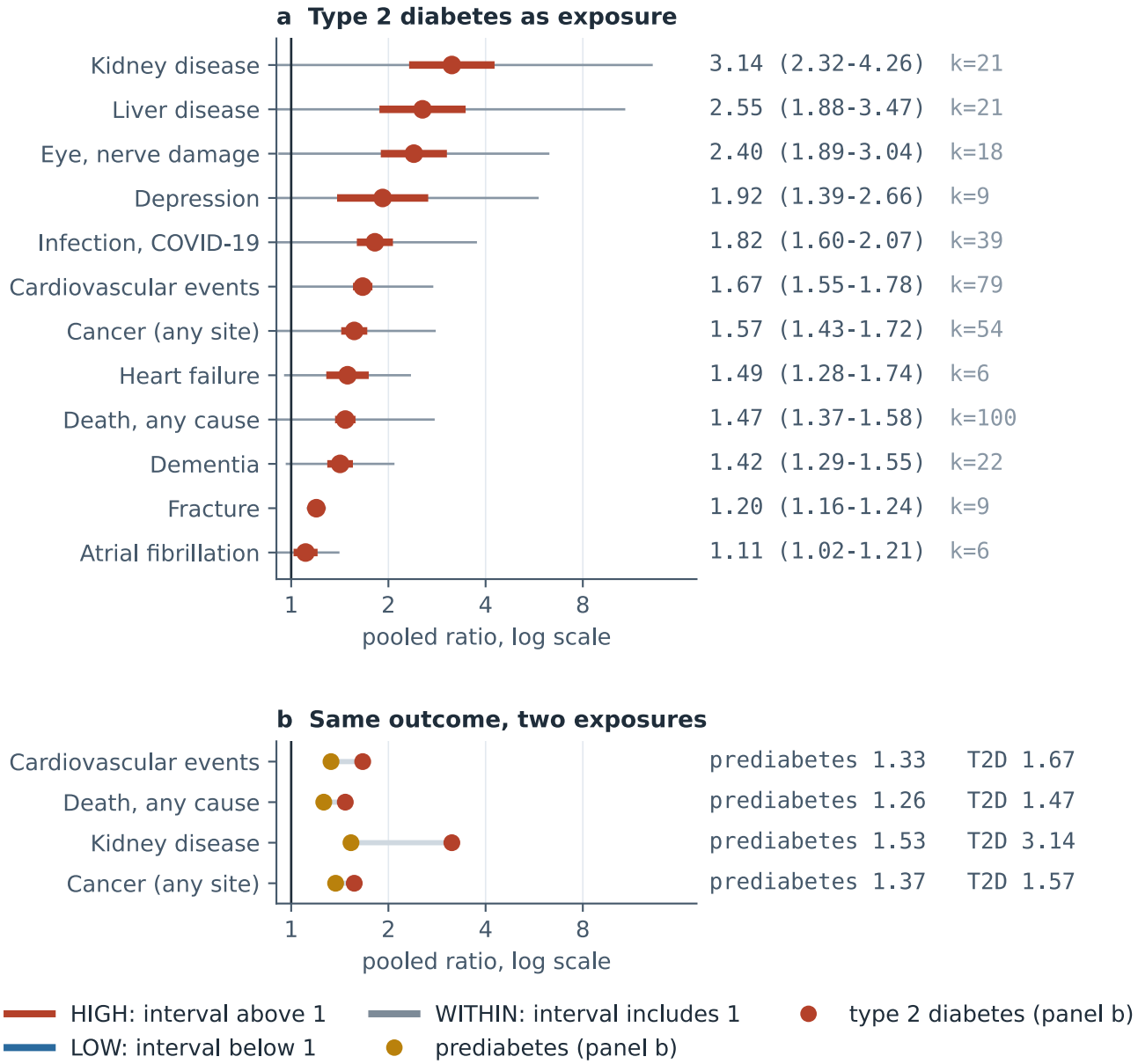


Figure 5. Panel C. (a) Type 2 diabetes as exposure, pooled ratio by outcome. (b) The same four outcomes with prediabetes (amber) and type 2 diabetes (rust) as exposure.

Kidney disease is the largest ratio (3.14 (2.32-4.26)) and the only one larger in the precise half (3.46 in the more precise half and 2.79 in the less precise half). Eye and nerve damage is 2.40 (1.89-3.04). Of the common outcomes, the precise half gives 1.49 for cardiovascular events, 1.31 for death and 1.31 for cancer, each below the all-study figure. Prediabetes sits below diabetes on all four shared outcomes (Figure 5b): cardiovascular events 1.33 against 1.67, death 1.26 against 1.47, kidney 1.53 against 3.14, cancer 1.37 against 1.57. That ordering fits a graded effect of glucose (The Emerging Risk Factors Collaboration, 2010; Huang et al., 2016; Tabák et al., 2012), but k is 4 to 14 for prediabetes and the cardiovascular ratio is gone in the precise half.

Within family, odds ratios are smaller than hazard or risk ratios by x0.83 (95% CI 0.76 to 0.91, $p < 0.001$), which survives the Bonferroni threshold. The direction runs against the usual expectation that odds ratios exceed risk ratios for common outcomes, and the sheet reports it as unexplained; comparison group and follow-up length are candidates. East Asian studies give x1.13 ($p = 0.044$), which does not survive.

> INTERPRETATION All twelve diabetes-as-exposure rows and all four prediabetes rows are outside range (HIGH), including atrial fibrillation at a small ratio. Of the ten diabetes-as-exposure rows that could be split, eight hold but shrink; kidney disease is the only one larger in the precise half. Prediabetes with cardiovascular events is gone in the precise half.

6 Glucose-lowering treatment

PANEL D

Two views, never pooled with each other. D1 is class-to-outcome ratios, the TRX lens. D2 is the effect on HbA1c and insulin resistance reported by systematic reviews, in the review's own units.

D1 Class to outcome

Specimen: 142 ratios from 112 studies, 139 observational. Pooled over all classes and outcomes the ratio is 0.77 (0.73-0.80), which mixes classes and is not a finding. Table 4 lists each class-outcome pair with at least four ratios.

Table 4. Panel D1. Glucose-lowering class and outcome. Ratio below 1.00 = fewer events. Observational ratios; comparison group varies between studies.

Class	Outcome	k	Ratio (95% CI)	95% PI	I²	Flag	Check
GLP-1 agonists	Death, any cause	5	0.65 (0.60 to 0.69)	0.56 to 0.75	5%	LOW	NOT SPLIT
GLP-1 agonists	Cardiovascular events	4	0.70 (0.59 to 0.82)	0.42 to 1.16	24%	LOW	NOT SPLIT
GLP-1 agonists	Cancer (any site)	6	0.97 (0.54 to 1.73)	0.13 to 7.36	93%	WITHIN	NOT SPLIT
GLP-1 agonists	Eye and nerve damage	4	1.03 (0.83 to 1.26)	0.41 to 2.56	87%	WITHIN	NOT SPLIT
Metformin	Death, any cause	14	0.79 (0.70 to 0.88)	0.51 to 1.21	96%	LOW	SMALLER
Metformin	Severe infection, COVID-19	4	0.82 (0.73 to 0.92)	0.53 to 1.25	50%	LOW	NOT SPLIT
Metformin	Cancer (any site)	12	0.88 (0.81 to 0.95)	0.69 to 1.11	68%	LOW	SMALLER
Metformin	Dementia and cognitive decline	6	1.08 (0.84 to 1.39)	0.47 to 2.51	98%	WITHIN	NOT SPLIT
SGLT2 inhibitors	Death, any cause	15	0.56 (0.44 to 0.72)	0.21 to 1.53	97%	LOW	HOLDS
SGLT2 inhibitors	Dementia and cognitive decline	5	0.66 (0.51 to 0.86)	0.25 to 1.77	97%	LOW	NOT SPLIT
SGLT2 inhibitors	Severe infection, COVID-19	4	0.70 (0.62 to 0.80)	0.42 to 1.18	68%	LOW	NOT SPLIT
SGLT2 inhibitors	Heart failure	5	0.73 (0.65 to 0.81)	0.52 to 1.01	50%	LOW	NOT SPLIT
SGLT2 inhibitors	Cardiovascular events	10	0.76 (0.62 to 0.94)	0.36 to 1.60	93%	LOW	HOLDS
SGLT2 inhibitors	Hypoglycaemia and diabetic ketoacidosis	5	2.27 (0.92 to 5.62)	0.08 to 61.51	92%	WITHIN	NOT SPLIT

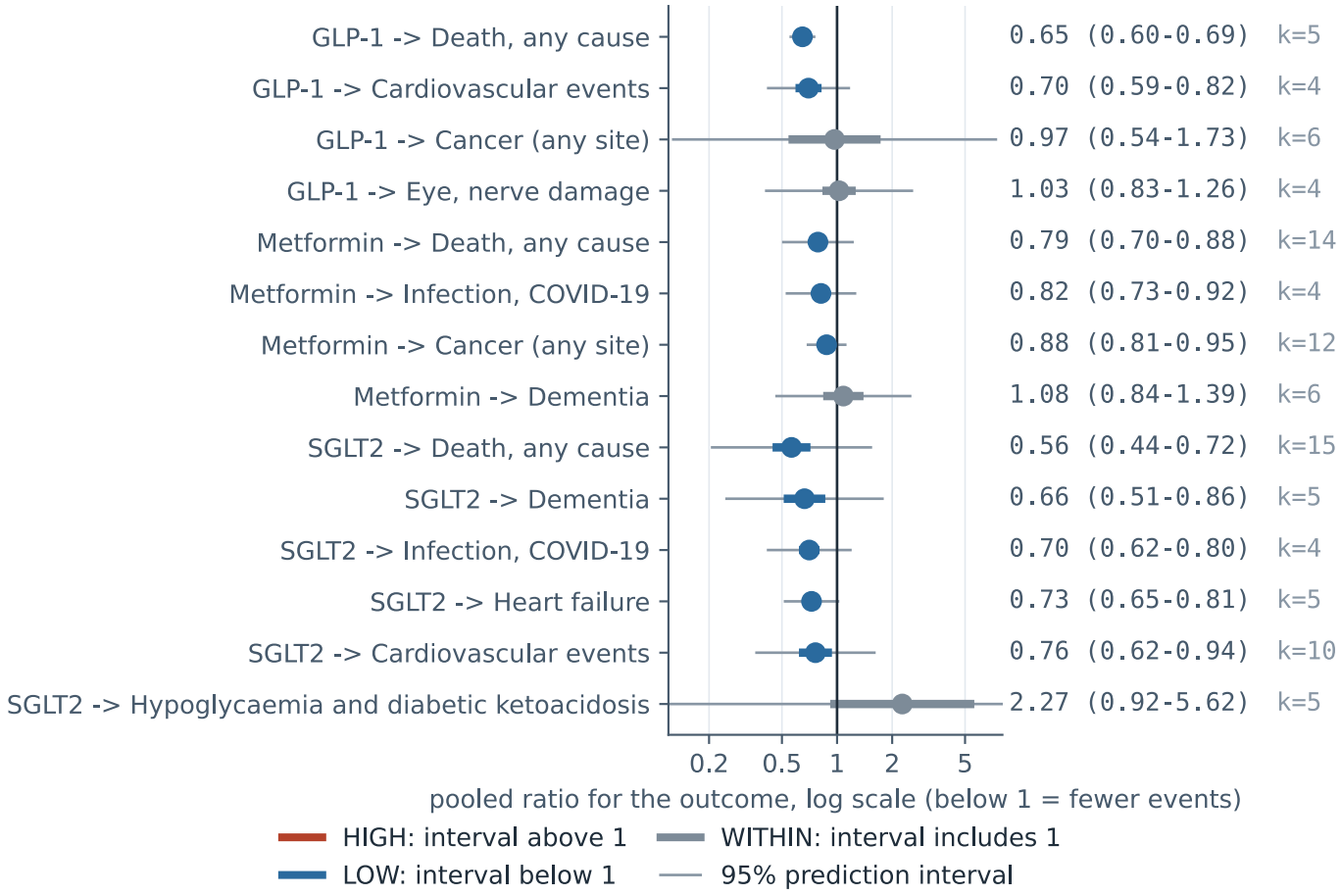


Figure 6. Panel D1. Class-to-outcome ratios; below 1.00 = fewer events. Bar: 95% CI; thin line: 95% prediction interval.

The rows are LOW for all-cause death in all three classes: SGLT2 0.56, GLP-1 0.65, metformin 0.79. The cardiovascular-outcome trials of these medicines (Zinman et al., 2015; Neal et al., 2017; Wiviott et al., 2019; Marso et al., 2016a; Marso et al., 2016b) and their pooled analyses (Sattar et al., 2021; Palmer et al., 2021) report hazard ratios for all-cause death nearer 1.00 (about 0.7 to 1.0) in people at high cardiovascular risk. For the two newer classes the observational ratios are further from 1.00 than those trials. For metformin an early trial in overweight people reported a comparable fall in deaths (Uk Prospective Diabetes

Study Group, 1998b). Two biases push the same way: immortal time between diagnosis and first prescription (Suissa, 2008), and healthier people being prescribed the newer drugs. Target-trial emulation addresses both (Hernán and Robins, 2016) and none of these ratios is shown to have used it. For SGLT2 and death the prediction interval runs from 0.21 to 1.53; a new study could find no benefit. Within lens, rows marked as adjusted give x0.78 (95% CI 0.64 to 0.96) against unadjusted, a pattern that does not survive Bonferroni.

Nulls: metformin and dementia is WITHIN (1.08 (0.84-1.39), k = 6), GLP-1 and cancer is WITHIN (0.97 (0.54-1.73), interval spans 0.54 to 1.73) and GLP-1 and eye or nerve damage is WITHIN (1.03 (0.83-1.26)). SGLT2 and hypoglycaemia or ketoacidosis (2.27 (0.92-5.62)) is the one outcome with a ratio above 1.00 and its interval includes it. The GLP-1 row for death has 5% I-squared on k = 5 and no halves; its narrow interval rests on five studies.

D2 What reviews report for HbA1c and insulin resistance

Specimen: 178 review-level estimates from 110 systematic reviews of people with diabetes or prediabetes, in the review's own scale. The pooled figure summarises what reviews report; reviews overlap in their trials, so it is not a new pooled trial estimate. Units are inferred from the sentence, not read from a field (Section 11).

Table 5. Panel D2. Mean differences reported by reviews, by intervention group (coarse groups only; fine classes failed the blind check). Negative = lower. Precise / imprecise = pooled difference in each half of the reviews.

Outcome	Group	k	Mean difference (95% CI)	95% PI	Precise / imprecise	Flag
HbA1c, percentage points	All reviews	49	-0.28 (-0.39 to -0.17)	-1.01 to 0.45	-0.20 / -0.37	LOW
HbA1c, percentage points	Lifestyle	23	-0.27 (-0.44 to -0.10)	-1.10 to 0.56	-0.15 / -0.41	LOW
HbA1c, percentage points	Supplements	11	-0.48 (-0.66 to -0.29)	-1.16 to 0.20	-0.42 / -0.52	LOW
HbA1c, percentage points	Drugs	10	-0.24 (-0.42 to -0.07)	-0.89 to 0.41	-0.15 / -0.35	LOW
HOMA-IR, mean difference	Lifestyle	4	-0.49 (-0.68 to -0.29)	-0.92 to -0.06	-	LOW
HOMA-IR, mean difference	Supplements	14	-0.63 (-0.77 to -0.49)	-0.99 to -0.27	-0.56 / -0.93	LOW
HOMA-IR, mean difference	Drugs	5	0.08 (-1.06 to 1.23)	-3.91 to 4.08	-	WITHIN

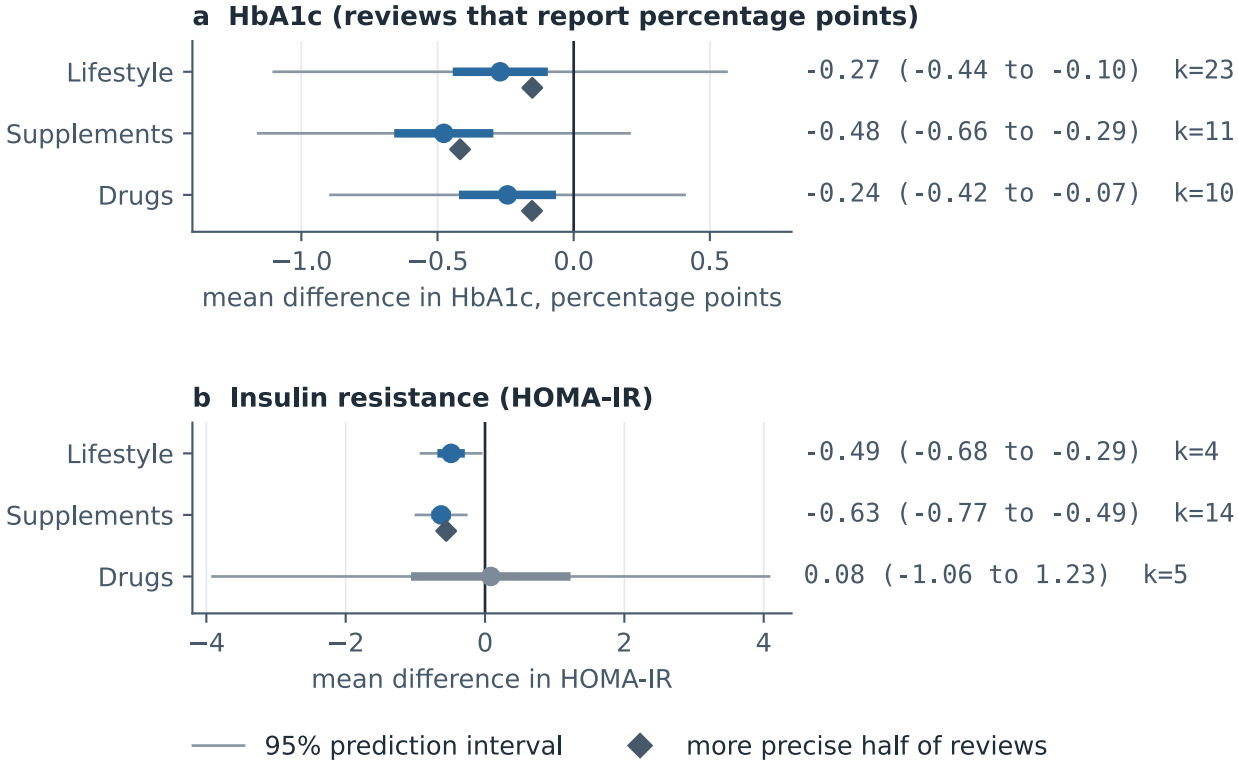


Figure 7. Panel D2. (a) HbA1c and (b) HOMA-IR, mean difference reported by reviews, by intervention group. Bar: 95% CI; thin line: 95% prediction interval; diamond: more precise half.

Reviews that report HbA1c in percentage points give -0.28 (95% CI -0.39 to -0.17). A fall of 0.5 percentage points is often used as a minimum worthwhile change, so the pooled review figure is below that and its prediction interval (-1.01 to 0.45) includes no change. The more precise half is -0.20 and the less precise half -0.37. The groups differ in what reviews choose to report: drug reviews often report add-on or sub-population comparisons, and supplement reviews pool small trials, so the larger supplement figure (-0.48) is not evidence that supplements beat drugs. The UK trials of intensive treatment against conventional treatment

gave HbA1c differences of about 0.6 to 0.9 points (Uk Prospective Diabetes Study Group, 1998a; Uk Prospective Diabetes Study Group, 1998b). HOMA-IR falls with lifestyle (-0.49, k = 4) and supplements (-0.63, k = 14); the drug row is WITHIN (0.08 (95% CI -1.06 to 1.23), k = 5) with I-squared 90%.

> INTERPRETATION Death rows for the three classes are outside range, low, and observational. Dementia and cancer rows for GLP-1 and metformin are within range. HbA1c from reviews is outside range but below the 0.5-point rule of thumb, and its prediction interval includes no change.

7 How common type 2 diabetes and prediabetes are

PANEL E

Specimen: 254 prevalence percentages with a denominator of at least 200 from 254 one-per-study rows, pooled on the logit scale, with prediction intervals. The diabetes row is a specimen of studies, not a national estimate. Two sources give population estimates directly (Sun et al., 2022; Ong et al., 2023; Zhou et al., 2016).

Table 6. Panel E. Pooled prevalence (logit scale). Median = median of study percentages.

Condition and group	k	Pooled % (95% CI)	95% PI	Median
Diabetes, general-population adults	127	10.3% (9.7 to 10.9)	5.4 to 18.9	10.2%
Prediabetes	40	20.0% (16.3 to 24.4)	4.5 to 57.3	19.9%
Undiagnosed diabetes	9	3.8% (2.8 to 5.1)	1.3 to 10.6	4.0%
Diabetes, older adults (60 and over)	14	16.4% (12.5 to 21.2)	4.8 to 43.2	16.4%
Diabetes, children and adolescents	6	4.5% (2.1 to 9.7)	0.2 to 49.3	3.3%
Diabetes, adults, North America	10	14.6% (11.2 to 18.7)	5.1 to 35.3	14.2%
Diabetes, adults, Latin America	6	14.1% (7.6 to 24.8)	1.2 to 69.0	13.4%
Diabetes, adults, Middle East & N. Africa	17	12.4% (10.4 to 14.7)	5.5 to 25.8	13.9%
Diabetes, adults, South & SE Asia	17	10.0% (7.7 to 12.9)	2.8 to 29.6	9.9%
Diabetes, adults, East Asia	24	8.7% (7.5 to 10.2)	3.8 to 18.9	10.2%
Diabetes, adults, Europe	12	8.7% (5.9 to 12.6)	1.7 to 34.2	9.2%
Diabetes, adults, Sub-Saharan Africa	17	7.2% (4.7 to 10.6)	1.1 to 35.7	7.4%

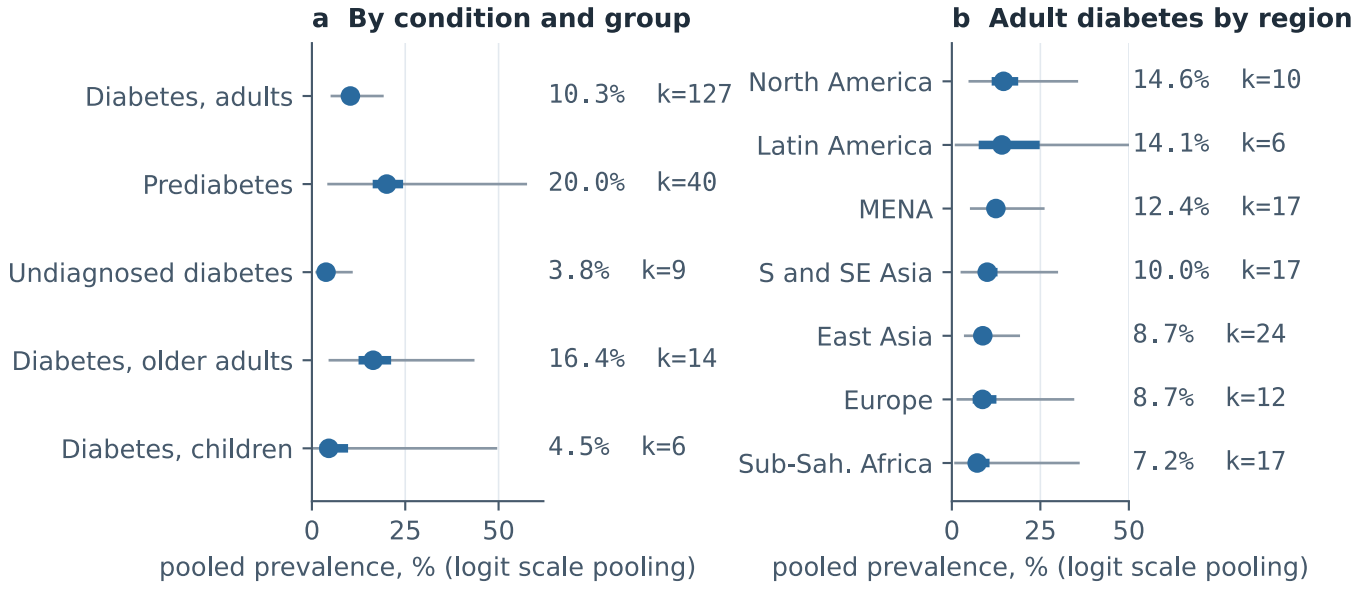


Figure 8. Panel E. Pooled prevalence with 95% CI (bar) and 95% prediction interval (thin line). (a) By condition and group; (b) adult diabetes by region.

General-population diabetes pools to 10.3% (95% CI 9.7 to 10.9), and the prediction interval is 5.4 to 18.9%: studies differ by a factor of three. Prediabetes pools to 20.0% (k = 40) with an interval of 4.5 to 57.3%; undiagnosed diabetes 3.8% (k = 9). By period the diabetes figure runs 8.9%, 10.9%, 11.8% with overlapping intervals; the typed layer is thin in 2024 and 2025, so a trend is not claimed. Regions differ less than the studies within them: North America against Europe is an odds ratio of 1.84 (p = 0.08). A 'definition' field (fasting glucose, HbA1c, self-report) failed the blind check and is not used.

Research context. The typed layer holds 35,903 papers that name diabetes or prediabetes, 48.7 per 10,000 typed papers in 2017 and 64.5 in 2024 (Figure 9). The countries named most often are China (1,313 papers), United States (1,186) and United Kingdom (742). These are context for the prevalence figures and not findings.

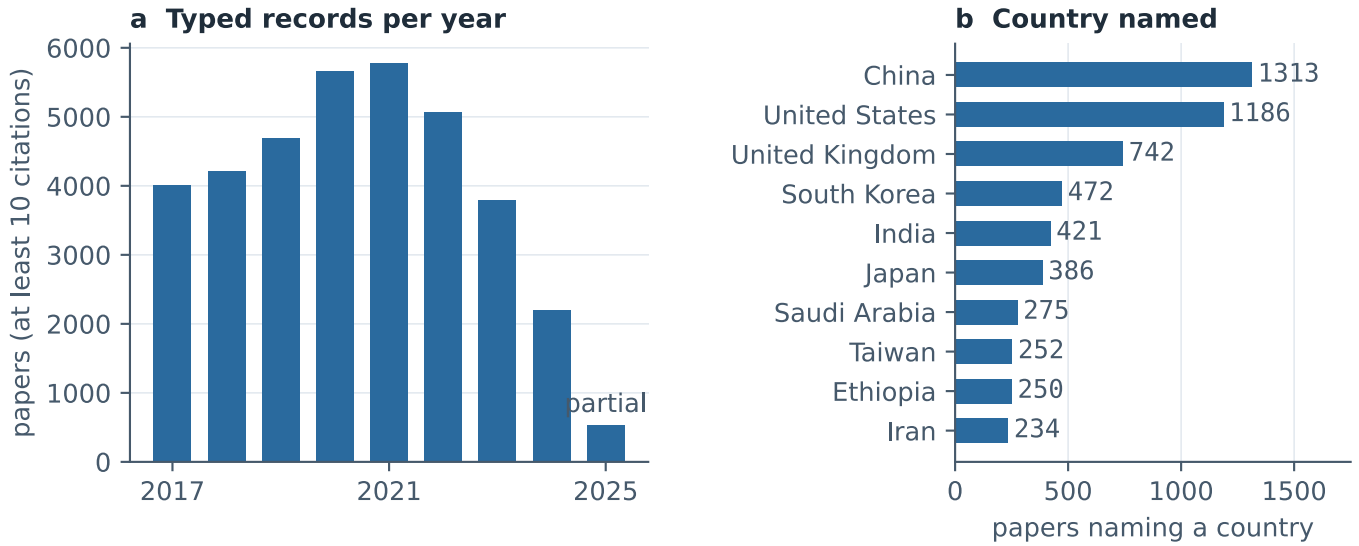


Figure 9. Context. (a) Typed research records per year that name diabetes or prediabetes (at least ten citations; 2025 is partial). (b) Countries named most often in those papers.

> INTERPRETATION Prevalence is outside any single reference range because studies disagree. The pooled figures describe the studies in the layer; the heterogeneity is the result.

8 What the news says

PANEL F

Specimen: 62,631 candidate articles matched a rule on the opening text (type 2 diabetes in its wordings, T2D, T2DM, NIDDM, prediabetes); 44,397 were removed as academic sources (by domain, then by a text classifier with precision 0.982 and recall 0.989), 105 as repeated boilerplate, leaving 18,129 articles. 9,160 of the 62,631 candidates have no publisher record because article-store nodes were unreachable. News is measured per 100,000 articles in the same month.

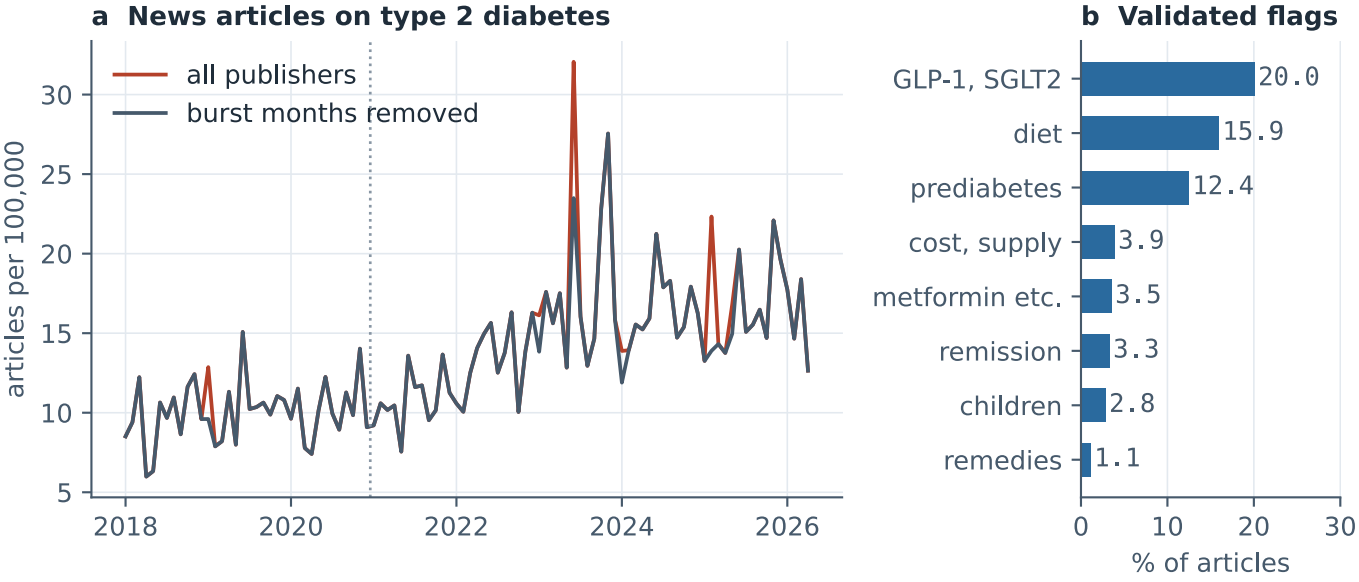


Figure 10. Panel F. (a) Articles on type 2 diabetes per 100,000 articles, monthly, 2018 to April 2026 (dotted line: start of the main window; the earlier months are the baseline extension). Red: all publishers; dark: months with a single-publisher burst removed. (b) Share of articles carrying each validated flag.

The rate is 10.0 per 100,000 in 2018-2019 and 16.9 from 2023 to April 2026, a ratio of 1.68 (1.64 without the burst months); the count ratio is 3.46 because the corpus itself grew 2.06 times. The rate peak, June 2023 (32.1 per 100,000), is one publisher, a biotech news site with 143 of 535 articles; the count peak, February 2025, is another publisher with 226 of 598. Neither is a public-interest event. One health-news site supplies 10.0% of articles with a publisher, and 1,799 publishers appear in all.

Table 7. Panel F. News subject flags: blind precision (share of flagged openings the second reader confirms), the round that decided the verdict, and the share of the corpus for flags that passed. Round 2 used rewritten rules on a fresh sample.

Flag	Round	n flagged	Precision	Verdict	Share of articles
Cites a study or report	1	169	86.4%	PASS	42.9%
GLP-1, SGLT2 and new drugs	1	82	98.8%	PASS	20.0%
Diet and foods	2	82	80.5%	PASS	15.9%
Prediabetes	1	60	93.3%	PASS	12.4%
Cost, prices and supply	1	30	76.7%	PASS	3.9%
Metformin and older drugs	1	30	93.3%	PASS	3.5%
Remission and reversal	2	38	94.7%	PASS	3.3%
Children and young people	2	37	86.5%	PASS	2.8%
Home remedies and supplements	2	34	85.3%	PASS	1.1%
Weight loss and exercise	2	87	44.8%	FAILED	not claimed
Complications	2	68	64.7%	FAILED	not claimed
How common it is	2	64	59.4%	FAILED	not claimed
Blood sugar tests and monitors	2	44	54.5%	FAILED	not claimed
Symptoms and signs	2	38	60.5%	FAILED	not claimed

The GLP-1 and SGLT2 flag climbs from 7.8% of articles in the first half of 2021 to 25.3% in the second half of 2025, with the step in the first half of 2023 (8.3% to 18.9%); metformin and older drugs stay near 3.5%. Diet and foods is 15.9%, prediabetes 12.4%, remission and reversal 3.3%, cost and supply 3.9% (rising to 7.5% in 2026), children 2.8%, home remedies 1.1%. 42.9% of openings cite a study or report. Five flags (weight and exercise, symptoms, blood sugar tests, complications, how common it is) failed twice and no claim is made about them.

Where news is placed: 52% of articles carry a country tag; of those the United States has 49.6%, the United Kingdom 11.1% and India 10.8%. The United States holds 12.5% of the research papers that name a country, a news-to-research ratio of 4.0; China has the largest research share (13.9%) against 3.7% of tagged news (Figure 12).

> INTERPRETATION The news corpus is outside range on volume (rate 1.68 times 2018-2019) and in its mix: drugs, diet and prediabetes dominate, while symptoms and complications are not claimed.

9 What people search for

PANEL G

Specimen: Google Trends, monthly relative interest 2004 to October 2026, 17 terms in joint and paired requests, two controls (weather, recipes), and 2,338 related-query pulls that return 618 distinct queries across six places and two windows. Values are scaled 0 to 100 inside each request, so only series from the same request are compared; every name below is also expressed as a ratio to the anchor ('type 2 diabetes') in the same two-term request (Nuti et al., 2014; Carneiro and Mylonakis, 2009). Repeated pulls returned identical values (correlation 1.0), so sampling noise could not be measured.

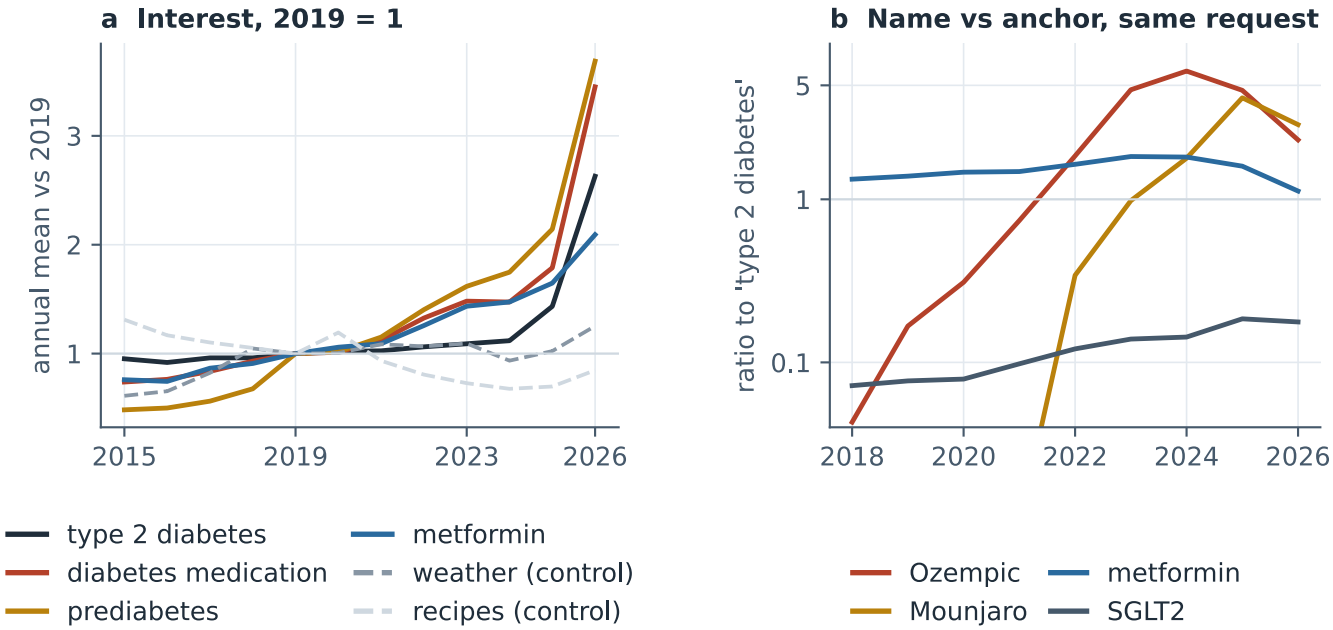


Figure 11. Panel G. (a) Annual mean search interest relative to 2019 (2026 is January to September). (b) Interest in four names relative to 'type 2 diabetes' in the same two-term request (log scale).

Against 2019, the anchor averages 1.10 in 2023-24 and 1.91 in 2025-26, while the weather control is 1.13 and recipes 0.75. In 2025-26: diabetes medication 2.46, prediabetes 2.76, metformin 1.82, SGLT2 4.2, HbA1c 2.46, Ozempic 37. Diabetes diet (1.14) and normal blood sugar level (0.98) stay near their 2019 level. The anchor itself rose, so every non-anchor series is given beside its ratio to the anchor (Figure 11b): Ozempic is 6.12 of the anchor in 2024 and 0.17 in 2019; Mounjaro peaks at 4.19 in 2025; metformin is stable at 1.39 to 1.82.

A 2026 surge sits on top: the anchor rose from 34 in January 2025 to 100 in May 2026, and March to June 2026 ran 2.5 to 2.9 times the median of the same 12-month span one to two years earlier. January to April 2026 against 2024-25 is 1.93 for the anchor, 1.85 for diabetes medication and 1.74 for prediabetes, against 1.19 for the weather control, 0.93 for Ozempic and 0.95 for the news rate. The surge is unexplained here: the news rate did not rise with it, and the control rose far less (the control's own 2026 level is also above its 2024-25 mean, so some of the lift may be tool-wide). It may be a measurement change in the tool, and the series ends with a partial month. Seasonality is weak: the November peak is 19% above the July trough.

Related queries, grouped by an intent rule that passed the blind check in 7 groups (Section 11), are led by medicines (33.8% of distinct queries) and tests and levels (24.3%). Then come the bare definition (11.5%), treatment and reversal (6.5%), symptoms (5.2%), diet and lifestyle (3.7%) and causes (1.9%). Complications are 0.3% (too few queries to validate). The 'other sense' group, which collects type 1 and non-medical queries, failed the check (20% agreement) and is not claimed. Among rising related queries across the seed terms, Ozempic appears at +2,350% and Mounjaro as a breakout.

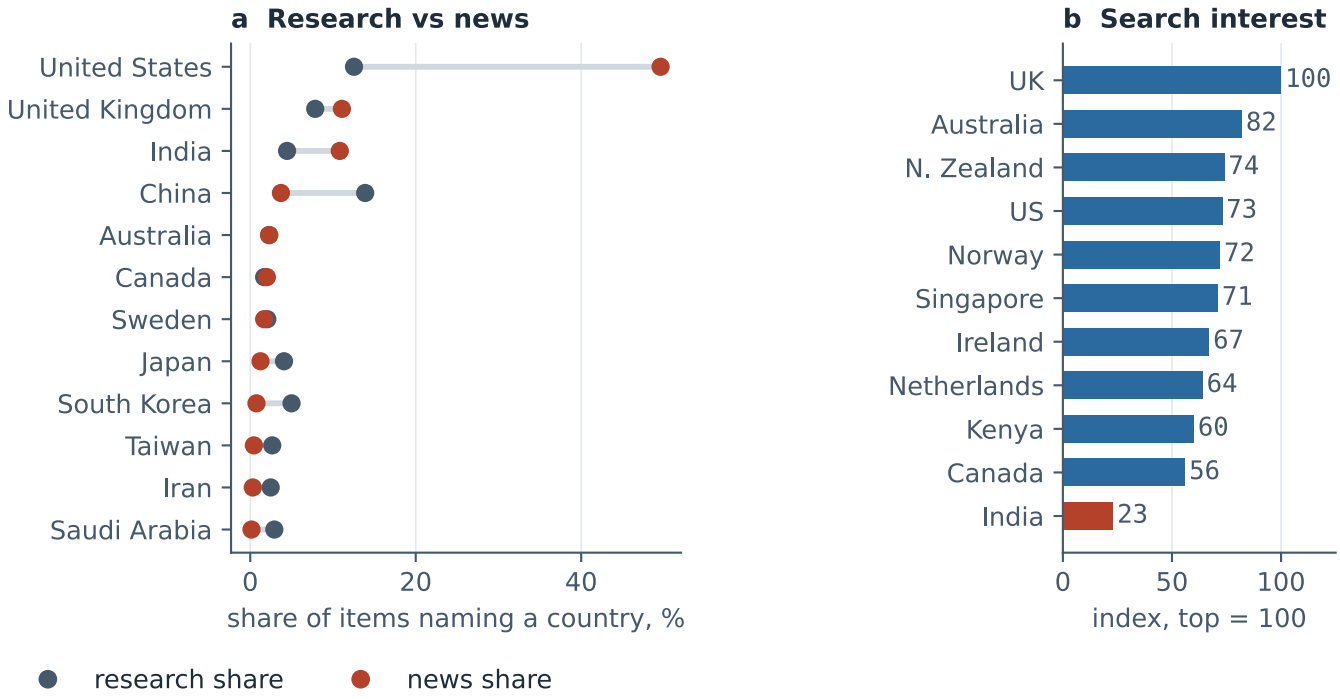


Figure 12. Panel G and F. (a) Country share of research papers naming a country (dark) and of news articles with a country tag (red). (b) Search interest in 'type 2 diabetes' by country in 2021 to 2026, top of the range = 100, India highlighted (small territories dropped).

By country, interest in the anchor is highest in the United Kingdom (index 100), then Australia (82), New Zealand (74) and the United States (73). India is 23. Within India the index runs from 100 in Puducherry to 18 in Bihar. India has 4.5% of the papers that name a country and 10.8% of tagged news; the national survey literature reports a large burden there (Anjana et al., 2023). Interest is relative inside each country, so countries are compared on rank, not volume.

> INTERPRETATION Search is outside range for 2025-26 (the anchor, medication and prediabetes terms all at 1.9 to 2.8 times 2019) with the control near range. Medicine and test queries dominate; symptom and diet queries are small.

10 Cross-source panel: research, news and search together

PANEL H

Two questions. Does attention follow the evidence? Does the monthly news rate move with search interest? For the second, both series are detrended and deseasonalised (linear trend plus month dummies), 100 months, January 2018 to April 2026. Only validated news measures enter.

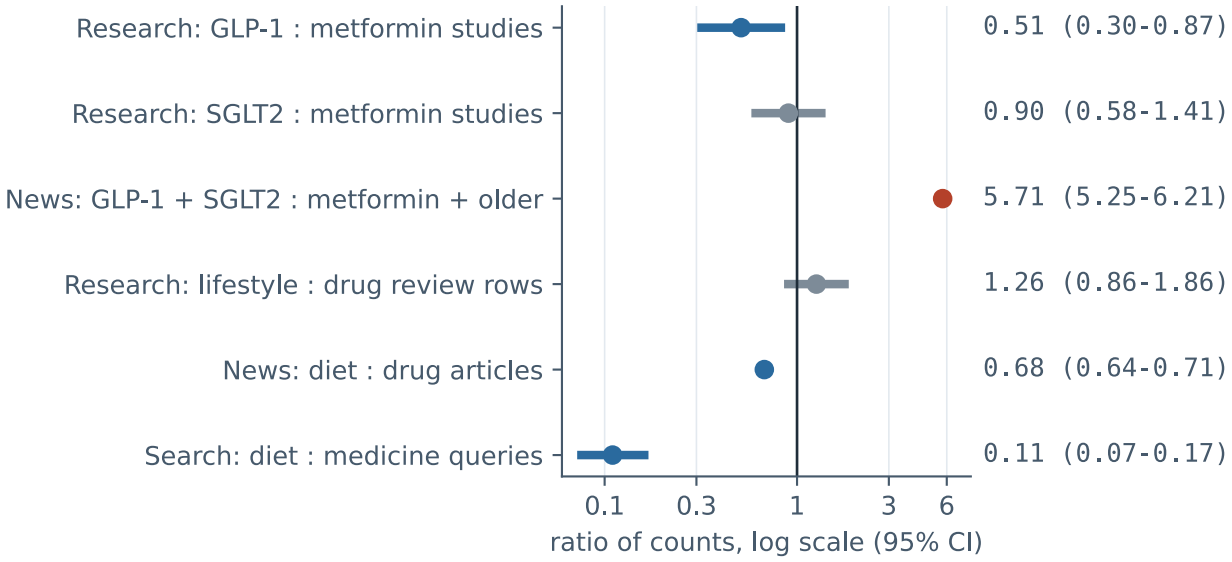


Figure 13. Panel H. Ratios of counts between source types, log scale (95% CI from counting error only; overlaps between flags and between studies are not removed).

In research the GLP-1 class has 0.51 times the studies of metformin (21 against 41; interval 0.30 to 0.87); SGLT2 inhibitors are close to metformin (37 against 41). In the news the new-drug flag is 5.7 times the metformin flag, and in search the brands are 4.9 times metformin. For lifestyle against drugs, the review rows split 58 to 46 (ratio 1.26, interval 0.86 to 1.86, includes 1.00);

news diet articles are 0.68 times drug articles; and diet queries are 0.11 times medicine queries. Lifestyle rows are about as common in the review layer as drug rows and are searched for far less than medicines. Recent papers have had less time to reach ten citations, so the typed layer understates the newest drug research, and the research ratio is a floor for GLP-1.

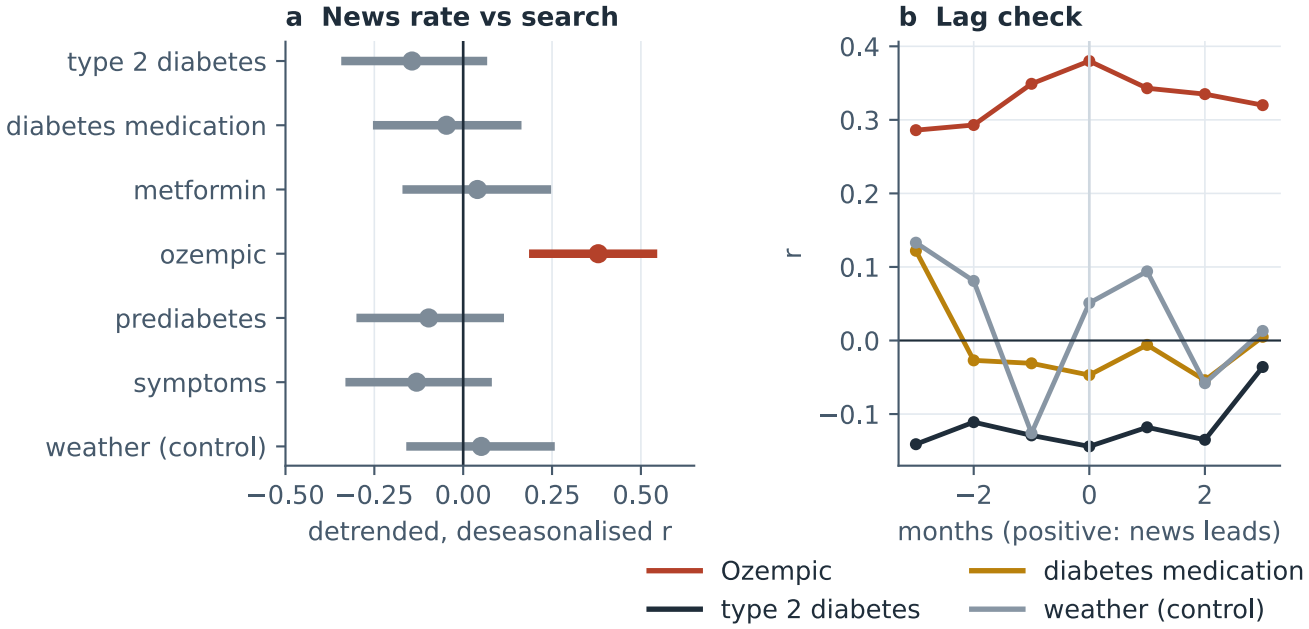


Figure 14. Panel H. (a) Detrended, deseasonalised correlation between the monthly news rate and monthly search interest, with 95% CI. (b) The same correlation with news shifted by -3 to +3 months (positive: news leads search).

The news rate and the anchor do not co-move after detrending (r = -0.14; metformin 0.04; prediabetes -0.10; weather control 0.05). The exception is Ozempic: r = 0.38 (95% CI 0.19 to 0.55), 0.38 inside 2021-2026 and 0.10 before. The lag curve peaks at zero, so news does not lead search. With publisher-burst months removed the figure is 0.34. The drug-name pair moves together in the news and in search, which is the only co-movement the sheet claims. Absence of positive co-movement for the other terms is not proof of independence; the two series are noisy.

> INTERPRETATION News and search weight the newer drugs several times more than metformin, while this research layer holds about half as many GLP-1 studies as metformin studies (a floor, because of the citation filter). Only Ozempic shows month-to-month co-movement between news and search.

11 Validation: blind second reader

QUALITY CONTROL

A fresh blind sample was drawn for every check. The second reader saw the extracted text and fields, never the pipeline label; truth files were sealed. Agreement of at least 70% is a pass, and a field that failed is not used in any claim. Rules were rewritten only between rounds and re-tested on fresh samples. The readers are AI readers (instances of Claude), not people; a second human reader was not available. Wilson intervals (Wilson, 1927).

Table 8. Validation summary. Agreement of the blind reader with the pipeline field (round of the verdict, sample size n). Full per-field bars in Figure 15.

Field or rule	Round	n	Agreement	Verdict
Ratio estimate and interval read from the quote	1	124	100%	pass
Role of diabetes in the row	1	124	88%	pass
Family of the other side	1	124	90%	pass
Risk population free of diabetes	1	54	100%	pass
Treatment class (first round: 54%)	2	44	77%	pass on re-test
Whole ratio row valid	1	124	73%	pass
Prevalence percentage, condition, general population	1	70	96% / 96% / 96%	pass
Prevalence definition stated	1	23	56%	FAILED, not used
Review outcome	1	64	89%	pass
Review HbA1c rows valid	1	22	91%	pass
Review unit read exactly	2	64	59%	FAILED; unit inferred, not contradicted in 95%
Review intervention group (coarse, first round: 61%)	2	64	81%	pass on re-test
Title-panel rules (risk, drugs, monitoring, complications, prevalence, lifestyle)	1	-	28% to 64%	FAILED, not used
Title-panel rules (prediabetes, young)	1	-	100% / 94%	pass (not otherwise used)
Search intent: medicines	1	30	97%	pass
Search intent: tests and levels	1	30	93%	pass
Search intent: bare definition	1	30	70%	pass
Search intent: treatment, remission, cure	1	30	93%	pass
Search intent: symptoms	1	30	97%	pass
Search intent: diet and lifestyle	1	23	100%	pass
Search intent: causes and risk groups	1	12	92%	pass
Search intent: other sense	1	30	20%	FAILED, not used
News flags (nine passed, five failed twice)	1 and 2	30 to 169	see Table 7	9 pass

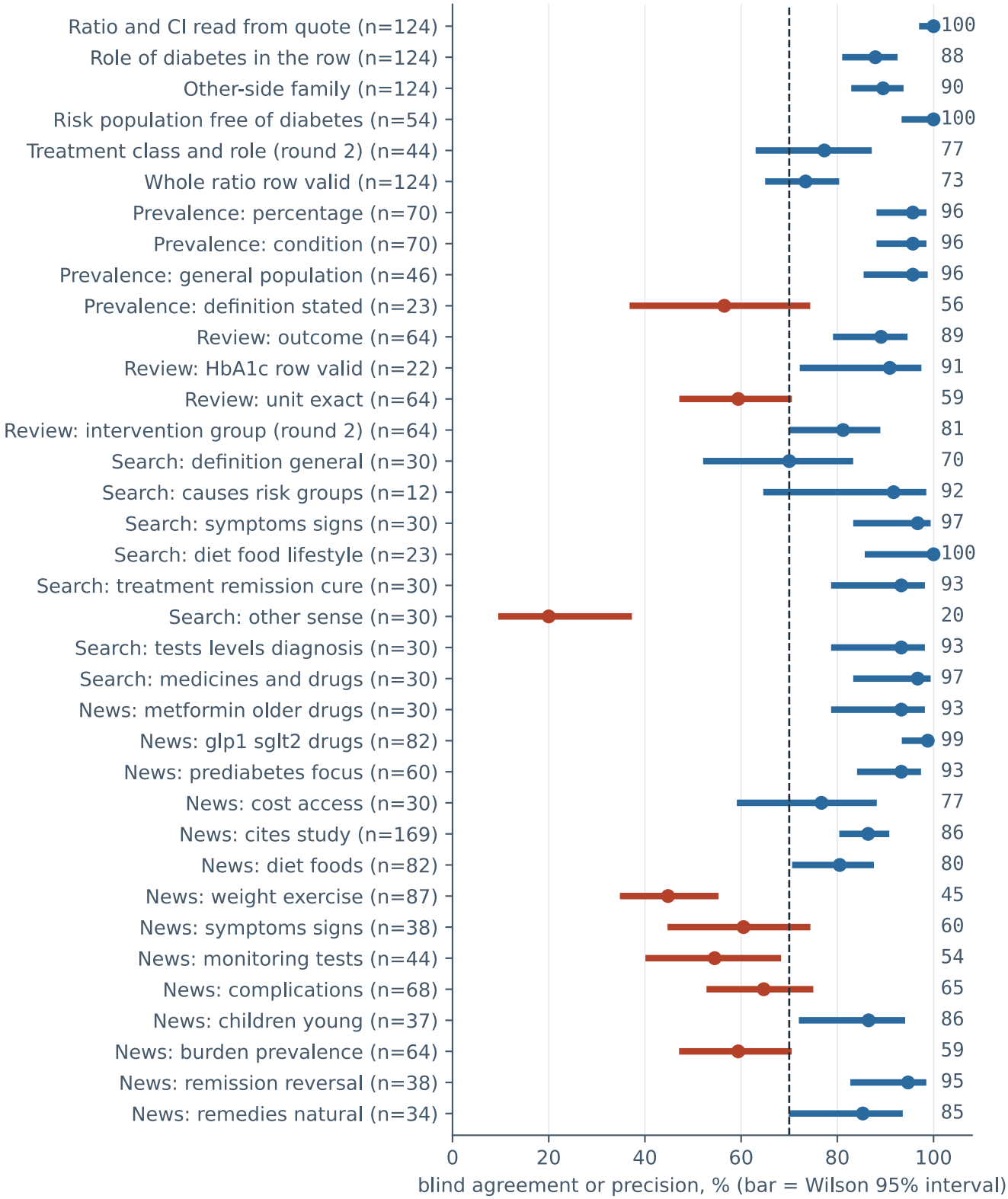


Figure 15. Quality control. Blind agreement or precision for every checked field, rule or flag, with Wilson 95% intervals; dashed line = 70% pass threshold.

News corpus precision, on random samples drawn under the headline rule: the reader judged 74.0% of 100 openings centrally about type 2 diabetes or prediabetes in round 1 and 65.5% of 110 in round 2; 100% and 99% named the condition at all and 0.0% in round 1 and 0.9% in round 2 used another sense. About a third of articles therefore mention the condition only in passing, and news volumes are counts of mentions, not of articles about the disease. Recall was low in the random strata for several flags, so shares are lower bounds.

12 Limits, and what is not claimed

NOTES

- **Observational ratios.** 478 of 485 risk ratios, 378 of 384 diabetes-as-exposure ratios and 139 of 142 treatment ratios are observational. Nothing here shows cause.
- **Study overlap.** One ratio per study, family and lens avoids counting a paper twice. Cohorts used by several papers, and trials used by several reviews, are not removed.
- **The typed layer.** Studies with at least ten citations only; thin in 2024 and 2025. Shares and counts for those years fall for that reason, not because research fell. Research counts come from six of seven shards (shard 5 was down), so counts are floors.
- **News coverage.** 13 of 14 news shards answered (shard 5 unreachable); 9,160 candidates lack a publisher record. News counts are mentions with a third incidental.
- **Search.** Relative interest only, one anchor that itself rose, identical replicate pulls, and a 2026 surge without a known cause.
- **Not claimed because the blind check failed:** prevalence by definition or criteria body; the exact unit of review estimates (the unit is inferred); fine intervention classes; title-based topic panels except prediabetes and young; the news flags for weight and exercise, symptoms, tests, complications and prevalence; the 'other sense' search group.
- **Not tested:** treatment effects on complications from trials (the layer holds no trial ratios in this lens); head-to-head drug comparisons; children, adolescents and pregnancy; type 1 and gestational diabetes.

13 Methods

PROCEDURE

Records. The extraction layer returns, per paper, findings with an exposure, an outcome, a quoted sentence and an effect size. 10,005 candidate ratios were re-read from the quoted sentence by a strict parser (ratio, lower and upper limit in one sentence; percentages refused). It removed 750: 147 with no number, 304 whose limits were not in one sentence, 229 with limits out of order and 70 given as percentages. Mapping then dropped 7,715 more, mostly exposures outside a listed family (3,251), animal, cell, type 1, pregnancy or COVID-only populations (1,258) and findings with no diabetes term (1,500); 84 risk ratios were removed because the population already had diabetes. 1,540 rows remained, 36 identical duplicates were removed, and one row per study, lens and family left 1,146; 95 per-unit rows were not pooled, leaving 1,051.

Pooling. Log ratios with standard errors from the interval, DerSimonian-Laird random effects, t prediction interval with k minus 2 degrees of freedom, precise and imprecise halves split at the median standard error within each family, within-family meta-regression with family dummies for imprecision and for each moderator, and a Bonferroni note for 20 moderator tests. Review mean differences pool on the linear scale, standardised and raw differences separately. Prevalence pools on the logit scale. No lens is pooled with another.

News. A rule on the opening text, an academic-source exclusion by domain and by a text classifier, a boilerplate exclusion, and fourteen subject flags with a proximity rule (the subject term within 100 characters of a diabetes term). **Search.** Google Trends through a local interface, joint batches and paired requests, controls, related queries with a first-match intent rule, and autocomplete. **Join.** Monthly series detrended and deseasonalised, lags of three months either way, within-regime correlations, and the weather control.

Table 9. Rows removed by the strict re-read and mapping (counts of ratio findings).

Step	Removed or kept	Count
Ratio candidates	kept for re-read	10,005
Strict re-read	removed	750
Mapping to a lens and family	removed	7,715
Identical duplicates	removed	36
One per study, lens and family	kept	1,146
Per-unit contrasts	not pooled	95
Pooled rows	analysed	1,051

14 Use of AI

This paper was not produced entirely by AI. The author chose the topic, set the objectives, the rules and the validation thresholds, reviewed every result and every decision, and is responsible for the text. AI (Claude, Anthropic) assisted with the analysis code, the figures, the drafting and the blind second reading. The blind second readers are AI readers (instances of Claude), not people. No result rests on a field that failed validation.

15 Competing interests and data

The records analysed come from the Fliar BI data platform (Prism); see references 1 and 2. Funding: none.

Derived tables (pooled results, validation scores, monthly series) are available from the author on request.

References

1. Anjana RM, Unnikrishnan R, Deepa M, Pradeepa R, Tandon N, Das AK, et al. (2023). Metabolic non-communicable disease health report of India: the ICMR-INDIAB national cross-sectional study (ICMR-INDIAB-17). *The Lancet Diabetes & Endocrinology* 11(7):474-489. https://doi.org/10.1016/S2213-8587(23)00119-5
2. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ (2015). Physical activity and the risk of type 2 diabetes: a systematic review and dose–response meta-analysis. *European Journal of Epidemiology* 30(7):529-542. https://doi.org/10.1007/s10654-015-0056-z
3. Carneiro H, Mylonakis E (2009). Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Diseases* 49(10):1557-1564. https://doi.org/10.1086/630200
4. DerSimonian R, Laird N (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* 7(3):177-188. https://doi.org/10.1016/0197-2456(86)90046-2
5. Egger M, Smith GD, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315(7109):629-634. https://doi.org/10.1136/bmj.315.7109.629
6. Hernán MA, Robins JM (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available: Table 1. *American Journal of Epidemiology* 183(8):758-764. https://doi.org/10.1093/aje/kwv254
7. Higgins JPT, Thompson SG, Spiegelhalter DJ (2008). A Re-Evaluation of Random-Effects Meta-Analysis. *Journal of the Royal Statistical Society Series A: Statistics in Society* 172(1):137-159. https://doi.org/10.1111/j.1467-985X.2008.00552.x
8. Huang Y, Cai X, Mai W, Li M, Hu Y (2016). Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. *BMJ* :i5953. https://doi.org/10.1136/bmj.i5953
9. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 6(7):e010247. https://doi.org/10.1136/bmjopen-2015-010247
10. Ioannidis JPA (2008). Why Most Discovered True Associations Are Inflated. *Epidemiology* 19(5):640-648. https://doi.org/10.1097/EDE.0b013e31818131e7
11. Malik VS, Popkin BM, Bray GA, Després JP, Willett WC, Hu FB (2010). Sugar-Sweetened Beverages and Risk of Metabolic Syndrome and Type 2 Diabetes. *Diabetes Care* 33(11):2477-2483. https://doi.org/10.2337/dc10-1079
12. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JF, Nauck MA, et al. (2016). Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine* 375(4):311-322. https://doi.org/10.1056/NEJMoa1603827
13. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, et al. (2016). Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *New England Journal of Medicine* 375(19):1834-1844. https://doi.org/10.1056/NEJMoa1607141
14. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. (2017). Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *New England Journal of Medicine* 377(7):644-657. https://doi.org/10.1056/NEJMoa1611925
15. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. (2014). The Use of Google Trends in Health Care Research: A Systematic Review. *PLoS ONE* 9(10):e109583. https://doi.org/10.1371/journal.pone.0109583
16. Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, et al. (2023). Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet* 402(10397):203-234. https://doi.org/10.1016/S0140-6736(23)01301-6
17. Palmer SC, Tendal B, Mustafa RA, Vandvik PO, Li S, Hao Q, et al. (2021). Sodium-glucose cotransporter protein-2 (SGLT-2) inhibitors and glucagon-like peptide-1 (GLP-1) receptor agonists for type 2 diabetes: systematic review and network meta-analysis of randomised controlled trials. *BMJ* :m4573. https://doi.org/10.1136/bmj.m4573
18. Sattar N, Lee MMY, Kristensen SL, Branch KRH, Del Prato S, Khurmi NS, et al. (2021). Cardiovascular, mortality, and kidney outcomes with GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis of randomised trials. *The Lancet Diabetes & Endocrinology* 9(10):653-662. https://doi.org/10.1016/S2213-8587(21)00203-5
19. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 343(jul22 1):d4002-d4002. https://doi.org/10.1136/bmj.d4002
20. Suissa S (2008). Immortal Time Bias in Pharmacoepidemiology. *American Journal of Epidemiology* 167(4):492-499. https://doi.org/10.1093/aje/kwm324
21. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. (2022). IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Research and Clinical Practice* 183:109119. https://doi.org/10.1016/j.diabres.2021.109119
22. Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M (2012). Prediabetes: a high-risk state for diabetes development. *The Lancet* 379(9833):2279-2290. https://doi.org/10.1016/S0140-6736(12)60283-9
23. The Emerging Risk Factors Collaboration (2010). Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *The Lancet* 375(9733):2215-2222. https://doi.org/10.1016/S0140-6736(10)60484-9
24. UK Prospective Diabetes Study (UKPDS) Group (1998). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *The Lancet* 352(9131):837-853. https://doi.org/10.1016/S0140-6736(98)07019-6
25. UK Prospective Diabetes Study (UKPDS) Group (1998). Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *The Lancet* 352(9131):854-865. https://doi.org/10.1016/S0140-6736(98)07037-8
26. Wilson EB (1927). Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association* 22(158):209-212. https://doi.org/10.1080/01621459.1927.10502953
27. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. (2019). Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine* 380(4):347-357. https://doi.org/10.1056/NEJMoa1812389
28. Zhou B, Lu Y, Hajifathalian K, Bentham J, Di Cesare M, Danaei G, et al. (2016). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. *The Lancet* 387(10027):1513-1530. https://doi.org/10.1016/S0140-6736(16)00618-8

29. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al. (2015). Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *New England Journal of Medicine* 373(22):2117-2128. https://doi.org/10.1056/NEJMoa1504720

Kumar N (2026). What the world studies, invents and reports about urban heat. Preprints.org. https://doi.org/10.20944/preprints202609.2322.v1

Fliar BI data platform (2026). Prism: the research, news and patent records used as the data source of this study. https://fliarbi.com

Fliar BI data platform (2026). Prism MCP: access tool for the research and news records [software, data source]. https://github.com/fliarbi/prism-mcp