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.

Markers of insulin resistance and six-year progression of anterior cervical and lumbar ossification: a Japanese community-based retrospective cohort study

Yu Li
Hirosaki University
Gentaro Kumagai
Hirosaki University
Sunao Tanaka
Hirosaki University
Kairo Wada
Hirosaki University
Takaaki Nakano
Hirosaki University
Sho Kaneko
Hirosaki University
Yasuyuki Ishibashi
Hirosaki University
Koichi Murashita
Hirosaki University
Kanichiro Wada
wadak39@hirosaki-u.ac.jp

Hirosaki University

Research Article

Keywords: Insulin resistance, Triglyceride–glucose index, HOMA-IR, Spinal ossification, Diffuse idiopathic skeletal hyperostosis, Longitudinal cohort

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11186784/v1

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

Additional Declarations: No competing interests reported.

Markers of insulin resistance and six-year progression of anterior cervical and lumbar ossification: a Japanese community-based retrospective cohort study

Abstract

Purpose To investigate the associations of baseline insulin resistance, assessed using the triglyceride–glucose (TyG) index and homeostasis model assessment of insulin resistance (HOMA-IR), with six-year anterior spinal ossification progression.

Methods This retrospective cohort included 307 XXX Health Promotion Project participants assessed in 2015 and 2021. Definite ossification increase (DOI) was evaluated across 13 cervical and lumbar segments. The principal exposures were TyG-defined insulin resistance (men ≥8.49; women ≥8.12) and continuous $\log_2$ HOMA-IR. Firth logistic regression incorporated demographic, adiposity, baseline ossification, and lifestyle covariates in successive models. Secondary outcomes were follow-up diffuse idiopathic skeletal hyperostosis (DISH) and spinal segmental ossification burden (SSOB).

Results DOI occurred in 37 participants. The proportions were 19.8% with and 8.3% without TyG-defined insulin resistance. The adjusted odds ratio was 2.29 (95% CI 1.05–5.11) after adjustment for age, sex, and waist circumference, and 2.44 (1.11–5.48) with additional lifestyle adjustment. Further adjustment for baseline SSOB attenuated the association to 1.64 (0.65–4.14). HOMA-IR was not associated with DOI; the corresponding fully adjusted odds ratio per doubling was 0.72 (0.39–1.30).

Conclusion TyG-defined insulin resistance was associated with six-year ossification progression after demographic, adiposity, and lifestyle adjustment. Attenuation after baseline SSOB adjustment suggests that this association overlaps with established skeletal burden. HOMA-IR showed no corresponding association.

Keywords: Insulin resistance; Triglyceride–glucose index; HOMA-IR; Spinal ossification; Diffuse idiopathic skeletal hyperostosis; Longitudinal cohort

Introduction

Diffuse idiopathic skeletal hyperostosis (DISH) is characterized by ligamentous and entheseal ossification, most prominently along the anterolateral spine [1]. Associations with obesity, diabetes, and metabolic syndrome have increasingly implicated metabolic dysfunction in its pathogenesis [2,3]. A previous cross-sectional study from the XXX Health Promotion Project identified an association between higher homeostasis model assessment of insulin resistance (HOMA-IR) and prevalent DISH [4]. These findings support a metabolic contribution to DISH, but longitudinal evidence linking insulin resistance to the development of spinal ossification remains limited.

Insulin resistance is a central feature of metabolic dysfunction and contributes to the development of dysglycemia, dyslipidemia, and related cardiometabolic disease [5,6]. Insulin signaling also regulates osteoblast differentiation and bone formation, providing a biological basis for an association between altered insulin action and skeletal hyperostosis [7]. Establishing whether insulin resistance contributes to progressive ossification would advance understanding of DISH etiology and could identify a modifiable metabolic pathway. This requires longitudinal assessment of metabolic status alongside structural changes in the spine.

This study examined the associations of baseline HOMA-IR and the triglyceride–glucose (TyG) index with six-year progression of anterior cervical and lumbar ossification in a Japanese community cohort. Follow-up DISH and total ossification burden were evaluated as secondary outcomes. We also assessed how these associations changed after accounting for adiposity, lifestyle, and baseline skeletal burden.

Methods

Study population

We hypothesized that greater baseline insulin resistance would be associated with subsequent anterior spinal ossification. Participants were drawn from the XXX Health Promotion Project, an annual community health-screening program in XXX, Japan [8]. This retrospective cohort included adults who attended in both 2015 and 2021. Among 328 participants, exclusions comprised unavailable target-segment scores, spinal surgery, ankylosing spondylitis, baseline exogenous insulin treatment, and long-term systemic glucocorticoid treatment. The resulting cohort included 307 participants (Fig. 1). Reporting followed the STROBE recommendations [9].

Assessment of insulin resistance

Baseline HOMA-IR was calculated as fasting glucose (mg/dL) × fasting insulin (μU/mL) / 405 [10]. A threshold of 2.5 has been used to identify insulin resistance in Japanese adults [11]. In this cohort, only two participants met that threshold, neither of whom developed DOI, limiting the precision of binary regression estimates. HOMA-IR was therefore analyzed primarily as a continuous variable. Its right-skewed distribution was $\log_2$-transformed to reduce the influence of extreme values and express associations per doubling. Supplementary

analyses compared the upper quartile with the lower three quartiles, following distribution-based approaches used in previous studies [4,12]. The cohort-specific 75th percentile (P75) indicated relative elevation.

TyG was calculated as ln[fasting triglycerides (mg/dL) × fasting glucose (mg/dL) / 2]. The primary TyG classification used sex-specific cut-offs of ≥8.49 in men and ≥8.12 in women, derived by Nakagomi et al. through receiver operating characteristic analysis using HOMA-IR ≥2.5 as the reference definition of IR [11]. Participants were classified as having TyG-defined IR or non-IR according to these cut-offs, which identify elevated HOMA-IR rather than establish a direct diagnosis of IR. Sensitivity analyses used a uniform threshold of 8.67, corresponding to the classic clamp-derived threshold after accounting for the formula convention [13,14], and the cohort P75. The latter approach has also been used in a large population-based cohort [15]. Continuous TyG, expressed per standard deviation (SD), was examined additionally. Cohort P75 values and the TyG SD were held constant across sensitivity analyses (Supplementary Methods S1).

Radiographic assessment

Anterior ossification was assessed on lateral radiographs using a study-specific ossification score (OS), informed by previous descriptive systems [16,17]. OS0 indicated no apparent ossification; OS1, anterior margin ossification without intervertebral bridging; OS2, ossification extending across the adjacent disc space without a complete bony bridge; and OS3, a complete bridge between adjacent vertebral bodies. Thirteen segments were assessed at both examinations: C1/2–C7/T1 and T12/L1–L5/S1. Two readers independently evaluated radiographs while blinded to image year and metabolic status, with third-reader adjudication.

No ossification increase (NOI) indicated no segmental increase. Equivocal ossification increase (EOI) comprised increases limited to OS0→OS1. Equivocal ossification increase (EOI) was assigned when the progression was limited to a one-grade increase in OS from baseline to follow-up but not reached OS3, as such a minimal change is difficult to reliably distinguish on lateral radiographs—even with two independent readers—and may be confounded by positional differences affecting the appearance of anterior margin ossification. Definite ossification increase (DOI) was assigned when progression reached OS3 at follow-up or two-grade increase in OS from baseline to follow-up, representing disc-space-spanning or complete bridge-forming ossification (Figure 2A-F). Unchanged OS3 was not classified as progression. The primary outcome was DOI versus non-DOI (NOI and EOI). Spinal segmental ossification burden (SSOB) was the sum of the 13 segmental scores, ranging from 0 to 39 (Fig. S1). DISH was diagnosed from the 2021 whole-spine lateral radiograph [1,4]. Follow-up DISH status and 2021 SSOB were secondary outcomes.

Covariates and adjustment strategy

Directed acyclic graphs guided the selection and interpretation of covariates [18,19]. Age, sex, central adiposity, and lifestyle were considered potential common determinants of metabolic status and skeletal progression. Waist circumference represented central adiposity; smoking and alcohol consumption were classified as never, former,

or current use, and exercise was assessed using the baseline nonwinter questionnaire. All covariates, including sex, were taken from 2015.

Baseline SSOB was considered separately because it represents established skeletal disease and may influence activity and subsequent progression, while also reflecting accumulated metabolic exposure. Models with and without baseline SSOB therefore addressed complementary questions: the overall longitudinal marker association and the association conditional on existing skeletal burden. Alternative temporal pathways are shown in Fig. 3 and detailed in Supplementary Methods S1.

For both exposures, M0 was unadjusted; M1 included age and sex; M2 added waist circumference; M3 added baseline SSOB; and M4 additionally included smoking, alcohol, and exercise. A parallel lifestyle model, M2-L, added these lifestyle variables to M2 without baseline SSOB. Comparison of M2-L with M4 evaluated the influence of conditioning on skeletal burden within an otherwise identical adjustment set.

Statistical analysis

Inter-reader agreement was assessed using two-way mixed-effects, single-measurement, absolute-agreement intraclass correlation coefficients [ICC(A,1)] for SSOB and its six-year change, linearly weighted Cohen's κ for segmental scores, and unweighted κ for progression classifications. The 95% CIs were obtained from 5,000 participant-level bootstrap resamples, retaining all segments within each participant.

Continuous variables were summarized as means (SDs) or medians (interquartile ranges), and categorical variables as counts and percentages. Standardized differences described baseline characteristics by TyG classification. Firth penalized logistic regression was used for DOI and DISH because of the limited event counts. Odds ratios (ORs), profile penalized-likelihood 95% confidence intervals (CIs), and penalized likelihood-ratio P values were calculated. Linear regression with HC3 robust standard errors and t-based inference estimated mean differences in follow-up SSOB. Continuous covariates entered as linear terms. TyG and HOMA-IR were fitted in separate models. Exposure, outcome, and model covariate data were complete for all 307 participants.

All six model specifications were reported with nominal two-sided P values. Supplementary Holm adjustments covered the two principal exposure–DOI tests within each specification and, separately, the 12 tests across all six specifications (Table S1). Benjamini–Hochberg adjustment covered 24 secondary tests comprising two exposures, two outcomes, and six specifications (Table S2). Sensitivity analyses evaluated alternative exposure definitions and excluded participants with recorded renal disease or urinary albumin-to-creatinine ratio ≥300 mg/g creatinine. Cohort size was determined by the available eligible participants. Analyses were performed in Python 3.12.14 and verified using R 4.4.1 with logistf and sandwich. Complete model estimates are provided in Tables S3–S12.

Results

Cohort characteristics

The 307 participants included 178 women (58.0%); mean baseline age was 51.2 years (SD 13.5). Over six years, 207 participants (67.4%) were classified as NOI, 63 (20.5%) as EOI, and 37 (12.1%) as DOI. Thirty participants (9.8%) had DISH at follow-up. Compared with the TyG-defined non-IR group, the IR group was older and had greater waist circumference and higher fasting triglyceride and insulin concentrations (Table 1).

DOI occurred in 20 of 101 participants (19.8%) with TyG-defined IR and 17 of 206 (8.3%) without it. The uniform TyG threshold of 8.67 identified 52 participants, including 13 DOI events. The TyG P75 was 8.45, identifying 77 participants and 19 events. Median HOMA-IR was 0.73 (interquartile range 0.54–1.08); its P75 identified 77 participants and 13 DOI events (Table 2). Exposure distributions are shown in Fig. S2.

Inter-reader agreement

Paired readings were available for all 307 participants. Total SSOB ICCs were 0.997 (95% CI 0.991–0.999) in 2015 and 0.985 (0.965–0.993) in 2021; the ICC for six-year change was 0.914 (0.844–0.953). Segmental scores agreed in 99.7% and 98.2% of readings, respectively (weighted κ, 0.982 and 0.918). DOI classification agreed in 297 participants (96.7%; κ=0.846, 95% CI 0.742–0.929). Both readers classified 37 participants as DOI, with 32 in common. Agreement for NOI/EOI/DOI was 92.8% (κ=0.852, 0.789–0.907) (Tables S13–S15).

Insulin resistance and ossification progression

TyG-defined IR was associated with DOI in the unadjusted model (OR 2.72, 95% CI 1.37–5.48), after age and sex adjustment (2.43, 1.16–5.19), and after additional waist adjustment (2.29, 1.05–5.11). The association remained with lifestyle adjustment in M2-L (2.44, 1.11–5.48). Estimates decreased after baseline SSOB was included: 1.62 (0.65–4.00) in M3 and 1.64 (0.65–4.14) in M4 (Table 3; Fig. 4). Nominal P values were 0.038 for M2 and 0.026 for M2-L; the corresponding within-model Holm-adjusted values were 0.076 and 0.053 (Table S1).

HOMA-IR was not significantly associated with DOI in any specification. ORs per doubling were 0.75 (0.45–1.26) in M2, 0.74 (0.43–1.28) in M3, and 0.72 (0.39–1.30) in M4. The HOMA-IR upper-quartile comparison yielded an OR of 0.92 (0.32–2.50) in M3. For continuous TyG, the M3 OR per SD was 0.97 (0.57–1.55). Alternative TyG thresholds and exclusion of 11 participants with renal conditions also produced CIs crossing unity after baseline SSOB adjustment (Tables S3–S12; Fig. S3).

Secondary outcomes

For follow-up DISH, M3 ORs were 1.42 (0.56–3.55) for TyG-defined IR and 0.77 (0.44–1.36) per HOMA-IR doubling. Corresponding mean differences in follow-up SSOB were 0.32 points (−0.06 to 0.70) and −0.17

points (−0.38 to 0.04). Results across all specifications are shown in Table 4. None of the 24 secondary tests remained significant at a Benjamini–Hochberg adjusted level of 0.05 (Table S2).

Discussion

In this six-year community cohort, TyG-defined IR was associated with progression of anterior spinal ossification after adjustment for age, sex, waist circumference, and lifestyle. This association attenuated when baseline ossification burden was included. HOMA-IR showed no corresponding longitudinal association. The findings extend the metabolic characterization of spinal ossification beyond cross-sectional DISH status and highlight the importance of existing skeletal disease when interpreting metabolic risk markers.

The association observed for TyG is consistent with the broader links between DISH and metabolic dysfunction [2,3]. The previous study related HOMA-IR to prevalent DISH [4], whereas the present study evaluated changes in individual cervical and lumbar segments over time. These outcomes capture different stages of the disease process: DISH reflects accumulated structural disease, while segmental progression includes changes below the threshold for a DISH diagnosis. The longitudinal TyG association therefore adds evidence that metabolic abnormalities accompany the evolution of anterior spinal ossification, including before extensive bridging develops.

Baseline SSOB materially influenced the estimates. TyG remained associated with DOI in M2-L, which included demographic, adiposity, and lifestyle factors, whereas adding SSOB in M4 reduced the OR from 2.44 to 1.64 and widened the uncertainty relative to unity. Existing ossification may identify local mechanical conditions and skeletal susceptibility that favor progression. It may also record the cumulative consequences of earlier metabolic abnormalities. Accordingly, adjustment for SSOB separates the marker association from established structural burden but may also remove part of a longer metabolic history. The attenuation is compatible with shared metabolic and skeletal influences; it does not quantify a mediated effect.

The differences between TyG and HOMA-IR may reflect the metabolic information captured by each index. TyG incorporates fasting triglyceride and glucose concentrations, whereas HOMA-IR depends on fasting insulin as well as glucose. Insulin signaling in osteoblasts provides biological plausibility for metabolic influences on bone formation [7], but circulating surrogate indices also reflect secretion, clearance, adiposity, and assay characteristics [20,21]. In this cohort, the low prevalence of HOMA-IR ≥2.5 and the narrow distribution of typical values limited discrimination by a conventional binary definition. Continuous modeling retained information across the distribution, although the resulting estimates remained imprecise.

Threshold selection also affected interpretation. The sex-specific TyG thresholds were derived against HOMA-IR in Japanese adults [11], providing population-relevant cut points for the principal analysis. The uniform threshold and upper-quartile analyses examined alternative definitions [13,14,15]. Binary contrasts and a continuous linear term describe different exposure patterns; the weaker continuous association indicates that the binary finding should not be generalized to a uniform increase in risk across the TyG range. Furthermore, the

nominal M2 and M2-L associations did not meet the supplementary two-exposure Holm criterion. These results support further evaluation of TyG as a metabolic marker of progression, with attention to threshold transportability and baseline skeletal burden.

Strengths include paired six-year radiographic assessment, quantified inter-reader agreement, adjudication of discrepant scores, segment-specific progression criteria, and parallel analyses of two insulin-resistance markers. Several limitations warrant consideration. The small number of DOI events limited precision, particularly in models containing multiple covariates. The community-screening population may differ from nonparticipants and patients with advanced symptomatic disease. Residual confounding by diet, medication, occupational loading, and lifetime metabolic exposure remains possible. Radiographs and ordinal grading may be insensitive to subtle structural changes. DISH was assessed at follow-up, so its association represents follow-up status rather than incident disease. Finally, the observational design and multiple analytical specifications limit causal interpretation. Larger cohorts with repeated metabolic and imaging assessments are needed to determine the temporal relationships between insulin resistance and ossification.

Conclusion

TyG-defined IR was associated with six-year anterior spinal ossification progression after demographic, adiposity, and lifestyle adjustment. Attenuation after adjustment for baseline SSOB suggests that the association overlaps with established skeletal burden. HOMA-IR showed no corresponding association. These findings support further investigation of metabolic dysfunction in spinal ossification progression.

References

1. Resnick D, Niwayama G (1976) Radiographic and pathologic features of spinal involvement in diffuse idiopathic skeletal hyperostosis (DISH). Radiology 119:559-568. https://doi.org/10.1148/119.3.559

2. Okada E, Ishihara S, Azuma K, Michikawa T, Suzuki S, Tsuji O et al (2021) Metabolic Syndrome is a Predisposing Factor for Diffuse Idiopathic Skeletal Hyperostosis. Neurospine 18:109-116. https://doi.org/10.14245/ns.2040350.175

3. Mader R, Novofestovski I, Adawi M, Lavi I (2009) Metabolic syndrome and cardiovascular risk in patients with diffuse idiopathic skeletal hyperostosis. Semin Arthritis Rheum 38:361-365. https://doi.org/10.1016/j.semarthrit.2008.01.010

4. Li Y, Wada K, Wada K, Nakano T, Kumagai G, Kaneko S, Koyama K, Tanaka S, Ishibashi Y (2026) Exploring the association between insulin resistance and diffuse idiopathic skeletal hyperostosis: a cross-sectional observational study. Spine J 26:942-956. https://doi.org/10.1016/j.spinee.2025.12.007

5. Reaven GM (1988) Banting lecture 1988. Role of insulin resistance in human disease. Diabetes 37:1595-1607. https://doi.org/10.2337/diab.37.12.1595

6. Samuel VT, Shulman GI (2016) The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. J Clin Invest 126:12-22. https://doi.org/10.1172/JCI77812

7. Fulzele K, Riddle RC, DiGirolamo DJ, Cao X, Wan C, Chen D et al (2010) Insulin receptor signaling in osteoblasts regulates postnatal bone acquisition and body composition. Cell 142:309-319. https://doi.org/10.1016/j.cell.2010.06.002

8. Nakaji S, Ihara K, Sawada K, Parodi S, Umeda T, Takahashi I, Murashita K, Kurauchi S, Tokuda I (2021) Social innovation for life expectancy extension utilizing a platform-centered system used in the Iwaki health promotion project: A protocol paper. SAGE Open Med 9:20503121211002606. https://doi.org/10.1177/20503121211002606

9. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 4:e296. https://doi.org/10.1371/journal.pmed.0040296

10. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC (1985) Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28:412-419. https://doi.org/10.1007/BF00280883

11. Nakagomi A, Sunami Y, Kawasaki Y, Fujisawa T, Kobayashi Y (2020) Sex difference in the association between surrogate markers of insulin resistance and arterial stiffness. J Diabetes Complications 34:107442. https://doi.org/10.1016/j.jdiacomp.2019.107442

12. Lee S, Choi S, Kim HJ, Chung YS, Lee KW, Lee HC, Huh KB, Kim DJ (2006) Cutoff values of surrogate measures of insulin resistance for metabolic syndrome in Korean non-diabetic adults. J Korean Med Sci 21:695-700. https://doi.org/10.3346/jkms.2006.21.4.695

13. Guerrero-Romero F, Simental-Mendía LE, González-Ortiz M, Martínez-Abundis E, Ramos-Zavala MG, Hernández-González SO, Jacques-Camarena O, Rodríguez-Morán M (2010) The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. J Clin Endocrinol Metab 95:3347-3351. https://doi.org/10.1210/jc.2010-0288

14. Simental-Mendía LE, Guerrero-Romero F (2020) The correct formula for the triglycerides and glucose index. Eur J Pediatr 179:1171. https://doi.org/10.1007/s00431-020-03644-1

15. Kazibwe R, Schaich CL, Chevli PA, Kingsley JA, Mirzai S, Namutebi JH et al (2026) Lipoprotein(a), Insulin Resistance, and Cardiovascular Disease in the UK Biobank. J Am Heart Assoc 15:e042361. https://doi.org/10.1161/jaha.125.042361

16. Mata S, Chhem RK, Fortin PR, Joseph L, Esdaile JM (1998) Comprehensive radiographic evaluation of diffuse idiopathic skeletal hyperostosis: development and interrater reliability of a scoring system. Semin Arthritis Rheum 28:88-96. https://doi.org/10.1016/S0049-0172(98)80041-3

17. Kuperus JS, Oudkerk SF, Foppen W, Mohamed Hoesein FA, Gielis WP, Waalwijk J et al (2019) Criteria for Early-Phase Diffuse Idiopathic Skeletal Hyperostosis: Development and Validation. Radiology 291:420-426. https://doi.org/10.1148/radiol.2019181695

18. Feeney T, Hartwig FP, Davies NM (2025) How to use directed acyclic graphs: guide for clinical researchers. BMJ 388:e078226. https://doi.org/10.1136/bmj-2023-078226

19. Greenland S, Pearl J, Robins JM (1999) Causal diagrams for epidemiologic research. Epidemiology 10:37-48. https://doi.org/10.1097/00001648-199901000-00005

20. Manley SE, Luzio SD, Stratton IM, Wallace TM, Clark PM (2008) Preanalytical, analytical, and computational factors affect homeostasis model assessment estimates. Diabetes Care 31:1877-1883. https://doi.org/10.2337/dc08-0097

21. Spoto B, Pisano A, Zoccali C (2016) Insulin resistance in chronic kidney disease: a systematic review. Am J Physiol Renal Physiol 311:F1087-F1108. https://doi.org/10.1152/ajprenal.00340.2016

Tables

Table 1 Baseline characteristics by TyG-defined insulin resistance

Characteristic	All N=307	TyG-defined non-IR N=206	TyG-defined IR N=101	SMD
Age, years	51.2 (13.5)	49.7 (14.1)	54.3 (11.6)	0.35
Female sex, n (%)	178 (58.0)	124 (60.2)	54 (53.5)	-0.14
BMI, kg/m²	22.2 (3.2)	21.5 (3.0)	23.8 (3.0)	0.75
Waist, cm	76.1 (9.7)	73.9 (9.2)	80.7 (9.1)	0.75
Fasting glucose, mg/dL	78.00 (72.50–85.00)	77.00 (72.00–82.00)	82.00 (76.00–89.00)	0.51
Triglycerides, mg/dL	75.00 (53.50–115.50)	60.00 (48.25–75.75)	140.00 (115.00–187.00)	1.27
Fasting insulin, μU/mL	3.80 (2.90–5.20)	3.50 (2.60–4.70)	4.80 (3.50–6.30)	0.61
TyG	8.10 (0.65)	7.76 (0.36)	8.80 (0.53)	2.28
HOMA-IR	0.73 (0.54–1.08)	0.65 (0.48–0.92)	0.93 (0.69–1.28)	0.60
Baseline total SSOB	0 (0–1)	0 (0–1)	0 (0–2)	0.29
Baseline cervical SSOB	0 (0–1)	0 (0–1)	0 (0–2)	0.33
Baseline lumbar SSOB	0 (0–0)	0 (0–0)	0 (0–0)	0.19
Lumbar SSOB >0, n (%)	48 (15.6)	25 (12.1)	23 (22.8)	0.28
Current smoking, n (%)	48 (15.6)	31 (15.0)	17 (16.8)	0.05
Former smoking, n (%)	75 (24.4)	49 (23.8)	26 (25.7)	0.05
Current alcohol use, n (%)	144 (46.9)	98 (47.6)	46 (45.5)	-0.04
Former alcohol use, n (%)	20 (6.5)	14 (6.8)	6 (5.9)	-0.04
Exercise, n (%)	89 (29.0)	52 (25.2)	37 (36.6)	0.25
Renal condition, n (%)	11 (3.6)	3 (1.5)	8 (7.9)	0.31

Values are mean (SD), median (Q1–Q3), or n (%). Glucose, triglycerides, insulin, HOMA-IR, and SSOB are shown as medians. All variables have N=307. SMD is the standardized difference (IR minus non-IR), calculated from means and SDs for continuous variables and proportions for binary variables. TyG-defined IR: men ≥8.49; women ≥8.12; non-IR: below these thresholds. Renal conditions: recorded renal disease or UACR ≥300 mg/g creatinine. IR, insulin resistance; TyG, triglyceride–glucose index; HOMA-IR, homeostasis model assessment of insulin resistance; SSOB, spinal segmental ossification burden; DOI, definite ossification increase; EOI, equivocal ossification increase; NOI, no ossification increase; DISH, diffuse idiopathic skeletal hyperostosis.

Table 2 Outcomes by insulin resistance classification and alternative exposure definitions

Exposure category	N	NOI, n (%)	EOI, n (%)	DOI, n (%)	DISH, n (%)
TyG-defined non-IR	206	145 (70.4)	44 (21.4)	17 (8.3)	13 (6.3)
TyG-defined IR	101	62 (61.4)	19 (18.8)	20 (19.8)	17 (16.8)
TyG <8.67	255	176 (69.0)	55 (21.6)	24 (9.4)	22 (8.6)
TyG ≥8.67	52	31 (59.6)	8 (15.4)	13 (25.0)	8 (15.4)
TyG <P75	230	165 (71.7)	47 (20.4)	18 (7.8)	16 (7.0)
TyG ≥P75	77	42 (54.5)	16 (20.8)	19 (24.7)	14 (18.2)
HOMA-IR <P75	230	157 (68.3)	49 (21.3)	24 (10.4)	17 (7.4)
HOMA-IR ≥P75	77	50 (64.9)	14 (18.2)	13 (16.9)	13 (16.9)

Percentages use the corresponding row denominator. TyG-defined IR: men ≥8.49; women ≥8.12. TyG P75=8.45 and HOMA-IR P75=1.08, rounded for display; group assignment used unrounded quantiles. P75 comparisons indicate relative elevation. Follow-up DISH was assessed in 2021. HOMA-IR ≥2.5 identified two participants with no DOI or DISH events. IR, insulin resistance; TyG, triglyceride–glucose index; HOMA-IR, homeostasis model assessment of insulin resistance; SSOB, spinal segmental ossification burden; DOI, definite ossification increase; EOI, equivocal ossification increase; NOI, no ossification increase; DISH, diffuse idiopathic skeletal hyperostosis.

Table 3 Principal exposures and definite ossification increase

Model	TyG-defined IR OR (95% CI)	P	HOMA-IR per doubling OR (95% CI)	P
M0	2.72 (1.37 to 5.48)	0.004	1.06 (0.69 to 1.65)	0.785
M1	2.43 (1.16 to 5.19)	0.019	0.90 (0.57 to 1.44)	0.646
M2	2.29 (1.05 to 5.11)	0.038	0.75 (0.45 to 1.26)	0.269
M3	1.62 (0.65 to 4.00)	0.292	0.74 (0.43 to 1.28)	0.272
M4	1.64 (0.65 to 4.14)	0.293	0.72 (0.39 to 1.30)	0.276
M2-L	2.44 (1.11 to 5.48)	0.026	0.72 (0.41 to 1.26)	0.244

N=307; DOI events=37 in every model. Firth logistic regression with profile-likelihood 95% confidence intervals (CIs) and nominal likelihood-ratio P values. TyG reference: TyG-defined non-IR. Multiplicity-adjusted P values appear in Table S1. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table 4 Principal exposures and secondary outcomes

Outcome	Model	TyG-defined IR Estimate (95% CI)	P	HOMA-IR per doubling Estimate (95% CI)	P
DISH OR	M0	2.97 (1.40 to 6.42)	0.005	1.45 (0.90 to 2.38)	0.130
DISH OR	M1	2.63 (1.19 to 5.95)	0.017	1.25 (0.76 to 2.13)	0.392
DISH OR	M2	1.86 (0.80 to 4.37)	0.148	0.82 (0.48 to 1.44)	0.473
DISH OR	M3	1.42 (0.56 to 3.55)	0.451	0.77 (0.44 to 1.36)	0.351
DISH OR	M4	1.38 (0.55 to 3.42)	0.488	0.84 (0.47 to 1.51)	0.553
DISH OR	M2-L	1.80 (0.78 to 4.21)	0.167	0.83 (0.47 to 1.49)	0.538
SSOB β	M0	1.53 (0.46 to 2.59)	0.005	0.36 (-0.11 to 0.84)	0.136
SSOB β	M1	1.02 (0.06 to 1.97)	0.037	0.11 (-0.27 to 0.49)	0.575
SSOB β	M2	0.90 (-0.04 to 1.83)	0.059	-0.09 (-0.52 to 0.35)	0.701
SSOB β	M3	0.32 (-0.06 to 0.70)	0.099	-0.17 (-0.38 to 0.04)	0.115
SSOB β	M4	0.32 (-0.07 to 0.71)	0.111	-0.17 (-0.39 to 0.05)	0.124
SSOB β	M2-L	0.90 (-0.06 to 1.87)	0.065	-0.13 (-0.56 to 0.30)	0.551

N=307; follow-up DISH cases=30. DISH estimates are Firth odds ratios. SSOB estimates are mean follow-up score differences (β, points) with HC3 confidence intervals and t-based inference. M3 and M4 include baseline SSOB. Nominal P values are shown; adjusted values appear in Table S2. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Figures

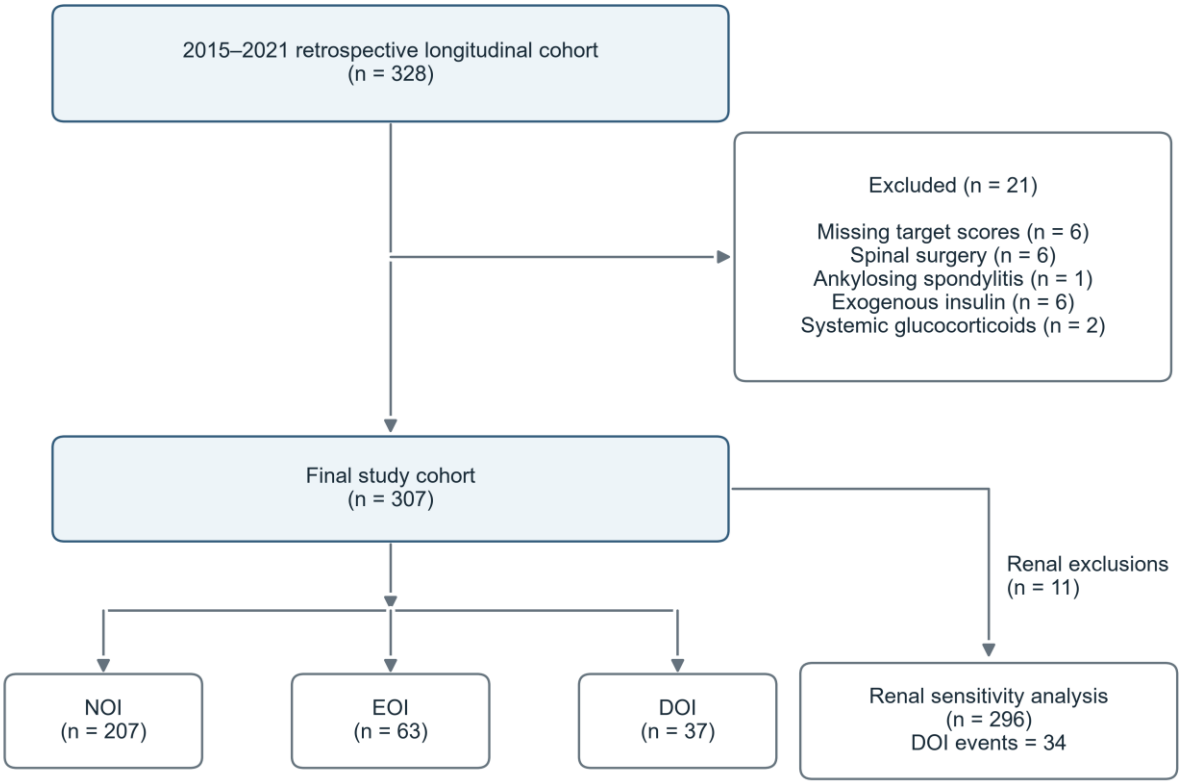


Fig. 1 Participant selection and analysis populations. Exclusion counts are mutually exclusive according to the displayed hierarchy. The renal sensitivity analysis excluded participants with recorded renal disease or urinary albumin-to-creatinine ratio ≥300 mg/g creatinine. NOI, no ossification increase; EOI, equivocal ossification increase; DOI, definite ossification increase

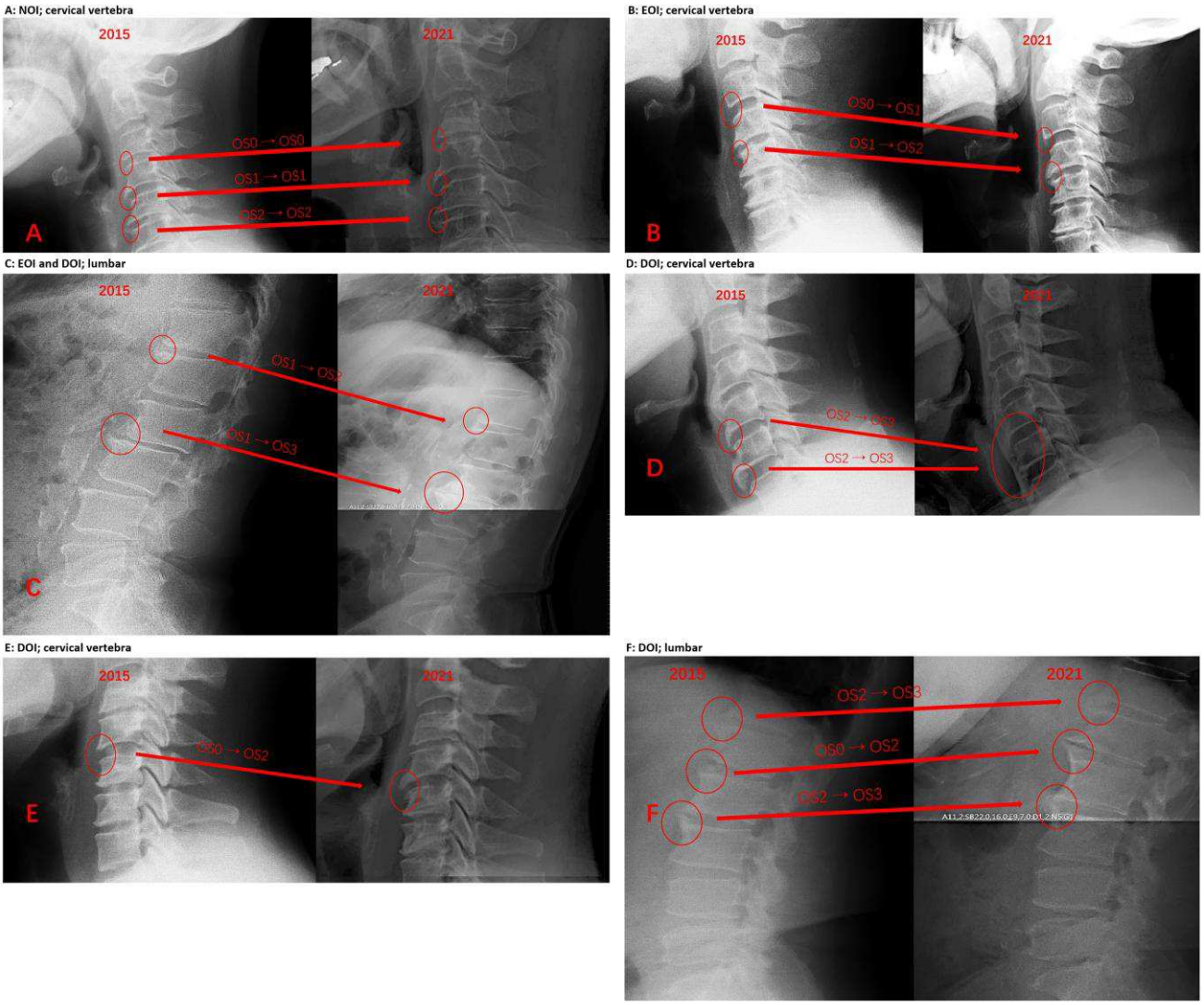


Fig. 2 Representative radiographic examples illustrating the imaging definitions used for NOI, EOI, and DOI between 2015 and 2021.

(A) shows NOI, with unchanged ossification findings at indexed cervical levels. (B) shows representative interval progression at cervical levels, including OS0-to-OS1 and OS1-to-OS2 change; under the study definitions, progression limited to OS1-level non-bridging ossification was classified as EOI. (C) shows lumbar examples of EOI and DOI, including a two-grade progression from OS0 to OS2 and progression to complete bridge formation (OS3). (D) shows cervical examples of DOI, including progression reaching OS2 and progression to complete bridge formation (OS3). Panels E and F show examples of DOI, including OS2-to-OS3 and OS0-to-OS2. Red circles and arrows identify indexed segments and the direction of change from baseline to follow-up.

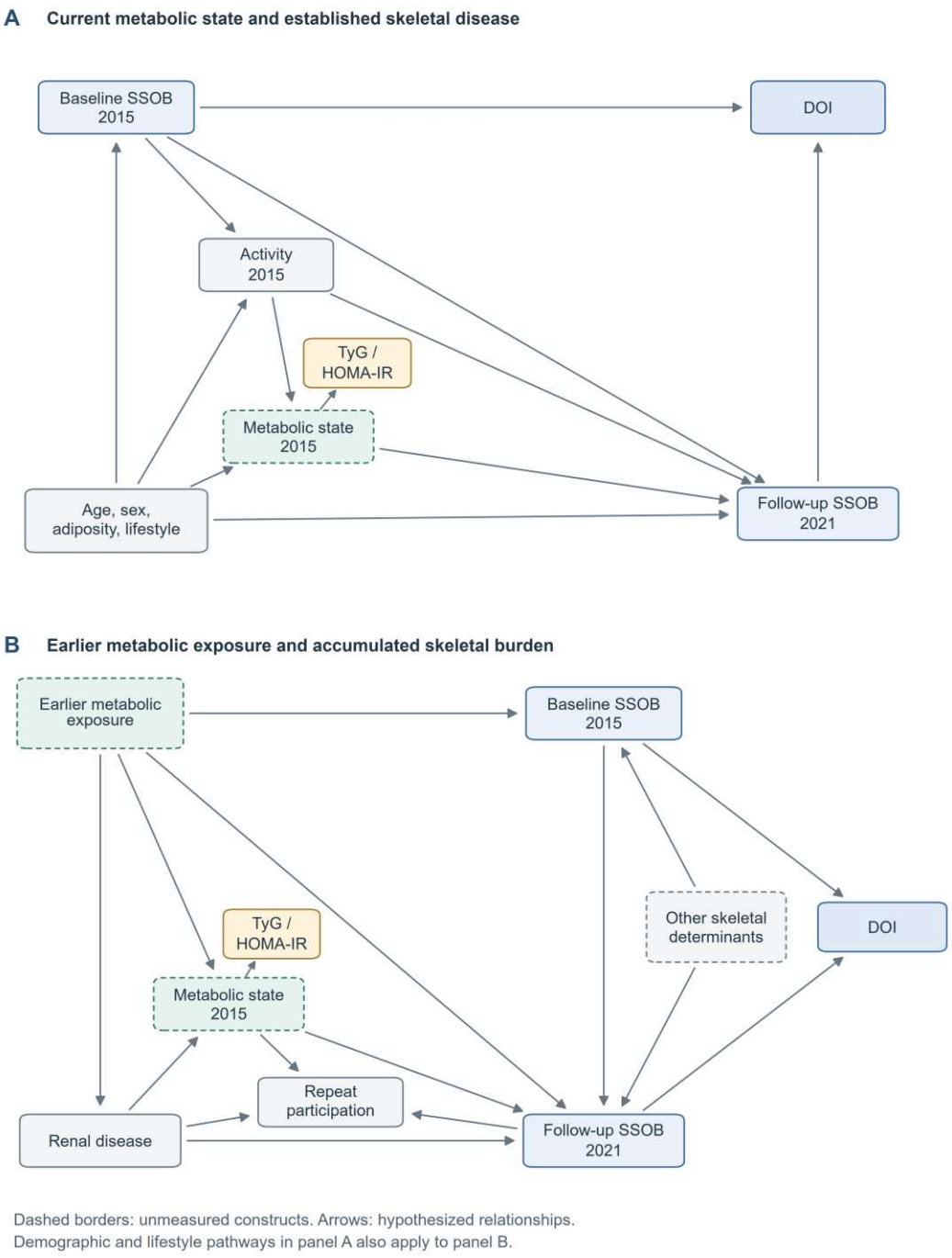


Fig. 3 Hypothesized relationships between metabolic state and spinal ossification. (A) Current metabolic state and established skeletal disease. (B) Earlier metabolic exposure and accumulated skeletal burden. Dashed borders denote unmeasured constructs. Demographic and lifestyle pathways in panel A also apply to panel B. Arrows into DOI indicate its dependence on baseline and follow-up segmental scoring; equal SSOB totals do not imply identical segmental patterns. SSOB, spinal segmental ossification burden; DOI, definite ossification increase.

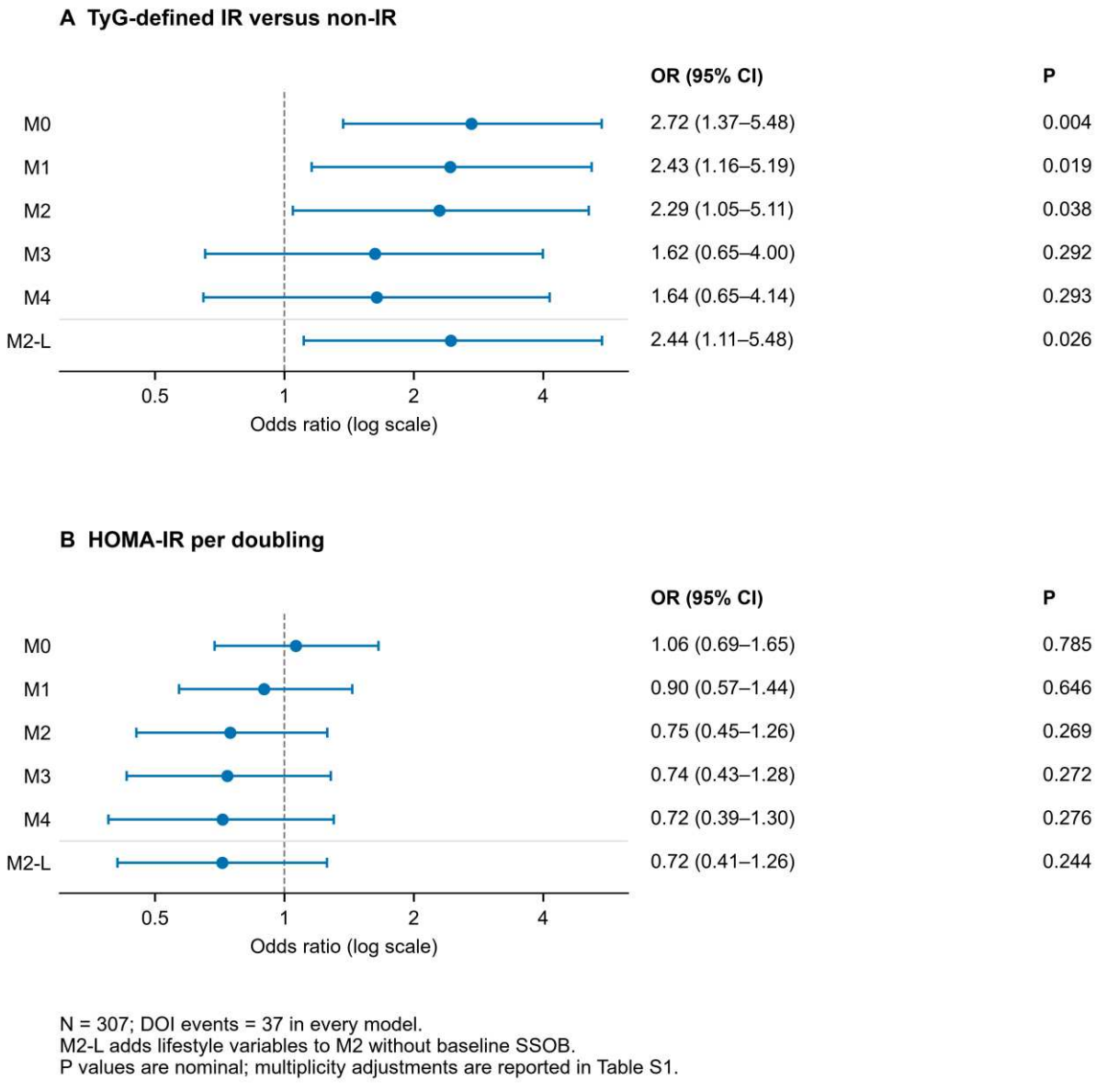


Fig. 4 Associations of principal insulin-resistance measures with DOI. (A) TyG-defined IR versus non-IR. (B) HOMA-IR per doubling. Points and horizontal lines indicate Firth odds ratios and profile-likelihood 95% confidence intervals. All models contain 307 participants and 37 DOI events. M2-L includes lifestyle covariates without baseline SSOB; M4 additionally includes baseline SSOB. P values are nominal

Supplementary Tables

Table S1 Multiplicity adjustments for the principal DOI hypotheses

Model	Exposure	Nominal P	Holm within level 2 tests	Holm across levels 12 tests
M0	TyG-defined IR	0.004	0.009	0.053
M0	HOMA-IR per doubling	0.785	0.785	1.000
M1	TyG-defined IR	0.019	0.038	0.210
M1	HOMA-IR per doubling	0.646	0.646	1.000
M2	TyG-defined IR	0.038	0.076	0.342
M2	HOMA-IR per doubling	0.269	0.269	1.000
M3	TyG-defined IR	0.292	0.545	1.000
M3	HOMA-IR per doubling	0.272	0.545	1.000
M4	TyG-defined IR	0.293	0.552	1.000
M4	HOMA-IR per doubling	0.276	0.552	1.000
M2-L	TyG-defined IR	0.026	0.053	0.265
M2-L	HOMA-IR per doubling	0.244	0.244	1.000

Holm adjustment was applied to the two principal exposure tests within each model and to all 12 principal DOI tests across models. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S2 Multiplicity adjustment for secondary outcomes

Outcome	Model	Exposure	Nominal P	BH adjusted 24 tests
DISH	M0	TyG-defined IR	0.005	0.062
DISH	M0	HOMA-IR per doubling	0.130	0.272
DISH	M1	TyG-defined IR	0.017	0.133
DISH	M1	HOMA-IR per doubling	0.392	0.589
DISH	M2	TyG-defined IR	0.148	0.273
DISH	M2	HOMA-IR per doubling	0.473	0.600
DISH	M3	TyG-defined IR	0.451	0.600
DISH	M3	HOMA-IR per doubling	0.351	0.561
DISH	M4	TyG-defined IR	0.488	0.600
DISH	M4	HOMA-IR per doubling	0.553	0.600
DISH	M2-L	TyG-defined IR	0.167	0.286
DISH	M2-L	HOMA-IR per doubling	0.538	0.600
SSOB21	M0	TyG-defined IR	0.005	0.062
SSOB21	M0	HOMA-IR per doubling	0.136	0.272
SSOB21	M1	TyG-defined IR	0.037	0.225
SSOB21	M1	HOMA-IR per doubling	0.575	0.600
SSOB21	M2	TyG-defined IR	0.059	0.262
SSOB21	M2	HOMA-IR per doubling	0.701	0.701
SSOB21	M3	TyG-defined IR	0.099	0.272
SSOB21	M3	HOMA-IR per doubling	0.115	0.272
SSOB21	M4	TyG-defined IR	0.111	0.272
SSOB21	M4	HOMA-IR per doubling	0.124	0.272
SSOB21	M2-L	TyG-defined IR	0.065	0.262
SSOB21	M2-L	HOMA-IR per doubling	0.551	0.600

Benjamini–Hochberg adjustment covers 24 tests: two principal exposures, two secondary outcomes, and six specifications. Effect estimates are reported in Table 4.

Table S3 TyG ≥8.67 in the full cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	307	37	3.23 (1.50 to 6.74)	0.003
DOI	M1	307	37	2.32 (1.01 to 5.18)	0.046
DOI	M2	307	37	2.14 (0.91 to 4.91)	0.080
DOI	M3	307	37	1.80 (0.61 to 4.93)	0.274
DOI	M4	307	37	1.88 (0.65 to 5.22)	0.239
DOI	M2-L	307	37	2.18 (0.93 to 5.02)	0.074
DISH	M0	307	30	1.98 (0.81 to 4.51)	0.131
DISH	M1	307	30	1.52 (0.59 to 3.66)	0.378
DISH	M2	307	30	1.00 (0.36 to 2.55)	0.995
DISH	M3	307	30	0.65 (0.19 to 1.89)	0.441
DISH	M4	307	30	0.65 (0.19 to 1.90)	0.450
DISH	M2-L	307	30	0.97 (0.35 to 2.47)	0.952
SSOB at 2021	M0	307	—	1.94 (0.23 to 3.66)	0.027
SSOB at 2021	M1	307	—	1.09 (-0.44 to 2.62)	0.163
SSOB at 2021	M2	307	—	0.94 (-0.56 to 2.44)	0.218
SSOB at 2021	M3	307	—	0.21 (-0.37 to 0.78)	0.481
SSOB at 2021	M4	307	—	0.21 (-0.37 to 0.79)	0.477
SSOB at 2021	M2-L	307	—	0.89 (-0.63 to 2.42)	0.251

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S4 TyG ≥P75 in the full cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	307	37	3.83 (1.90 to 7.76)	<0.001
DOI	M1	307	37	2.57 (1.21 to 5.49)	0.014
DOI	M2	307	37	2.44 (1.09 to 5.48)	0.030
DOI	M3	307	37	1.92 (0.76 to 4.81)	0.167
DOI	M4	307	37	1.99 (0.77 to 5.14)	0.154
DOI	M2-L	307	37	2.59 (1.16 to 5.88)	0.021
DISH	M0	307	30	2.97 (1.38 to 6.36)	0.006
DISH	M1	307	30	2.15 (0.95 to 4.82)	0.066
DISH	M2	307	30	1.43 (0.59 to 3.37)	0.420
DISH	M3	307	30	1.10 (0.41 to 2.78)	0.846
DISH	M4	307	30	1.03 (0.39 to 2.62)	0.944
DISH	M2-L	307	30	1.38 (0.57 to 3.24)	0.465
SSOB at 2021	M0	307	—	2.05 (0.75 to 3.36)	0.002
SSOB at 2021	M1	307	—	1.10 (-0.08 to 2.28)	0.067
SSOB at 2021	M2	307	—	0.97 (-0.20 to 2.14)	0.105
SSOB at 2021	M3	307	—	0.49 (0.03 to 0.95)	0.038
SSOB at 2021	M4	307	—	0.49 (0.01 to 0.97)	0.045
SSOB at 2021	M2-L	307	—	0.99 (-0.22 to 2.20)	0.109

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S5 TyG per SD in the full cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	307	37	1.59 (1.17 to 2.18)	0.003
DOI	M1	307	37	1.26 (0.87 to 1.81)	0.223
DOI	M2	307	37	1.19 (0.80 to 1.75)	0.389
DOI	M3	307	37	0.97 (0.57 to 1.55)	0.898
DOI	M4	307	37	1.04 (0.62 to 1.67)	0.875
DOI	M2-L	307	37	1.23 (0.82 to 1.82)	0.306
DISH	M0	307	30	1.56 (1.12 to 2.17)	0.009
DISH	M1	307	30	1.37 (0.91 to 2.03)	0.126
DISH	M2	307	30	1.11 (0.70 to 1.69)	0.653
DISH	M3	307	30	0.94 (0.55 to 1.52)	0.799
DISH	M4	307	30	0.93 (0.55 to 1.49)	0.766
DISH	M2-L	307	30	1.08 (0.69 to 1.66)	0.728
SSOB at 2021	M0	307	—	0.86 (0.36 to 1.37)	<0.001
SSOB at 2021	M1	307	—	0.25 (-0.18 to 0.68)	0.247
SSOB at 2021	M2	307	—	0.15 (-0.28 to 0.57)	0.498
SSOB at 2021	M3	307	—	-0.04 (-0.21 to 0.14)	0.670
SSOB at 2021	M4	307	—	-0.04 (-0.22 to 0.13)	0.619
SSOB at 2021	M2-L	307	—	0.15 (-0.29 to 0.59)	0.504

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S6 HOMA-IR ≥P75 in the full cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	307	37	1.76 (0.84 to 3.58)	0.131
DOI	M1	307	37	1.28 (0.57 to 2.76)	0.546
DOI	M2	307	37	1.07 (0.44 to 2.50)	0.884
DOI	M3	307	37	0.92 (0.32 to 2.50)	0.871
DOI	M4	307	37	1.09 (0.36 to 3.09)	0.877
DOI	M2-L	307	37	1.10 (0.44 to 2.64)	0.829
DISH	M0	307	30	2.55 (1.17 to 5.46)	0.019
DISH	M1	307	30	1.80 (0.79 to 4.03)	0.162
DISH	M2	307	30	1.06 (0.42 to 2.57)	0.900
DISH	M3	307	30	0.93 (0.35 to 2.40)	0.882
DISH	M4	307	30	1.02 (0.36 to 2.74)	0.973
DISH	M2-L	307	30	1.04 (0.39 to 2.64)	0.933
SSOB at 2021	M0	307	—	1.10 (-0.15 to 2.34)	0.084
SSOB at 2021	M1	307	—	0.46 (-0.64 to 1.57)	0.412
SSOB at 2021	M2	307	—	0.22 (-1.09 to 1.54)	0.738
SSOB at 2021	M3	307	—	-0.22 (-0.70 to 0.26)	0.360
SSOB at 2021	M4	307	—	-0.22 (-0.71 to 0.27)	0.380
SSOB at 2021	M2-L	307	—	0.12 (-1.15 to 1.39)	0.851

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S7 TyG-defined IR in the renal-restricted cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	296	34	2.44 (1.19 to 5.01)	0.015
DOI	M1	296	34	2.01 (0.93 to 4.36)	0.075
DOI	M2	296	34	1.91 (0.85 to 4.32)	0.117
DOI	M3	296	34	1.36 (0.53 to 3.44)	0.515
DOI	M4	296	34	1.32 (0.50 to 3.44)	0.574
DOI	M2-L	296	34	2.02 (0.89 to 4.61)	0.090
DISH	M0	296	29	3.00 (1.40 to 6.56)	0.005
DISH	M1	296	29	2.56 (1.14 to 5.86)	0.023
DISH	M2	296	29	1.86 (0.79 to 4.43)	0.156
DISH	M3	296	29	1.53 (0.60 to 3.85)	0.367
DISH	M4	296	29	1.50 (0.59 to 3.75)	0.386
DISH	M2-L	296	29	1.82 (0.77 to 4.32)	0.168
SSOB at 2021	M0	296	—	1.41 (0.37 to 2.44)	0.008
SSOB at 2021	M1	296	—	0.83 (-0.10 to 1.77)	0.082
SSOB at 2021	M2	296	—	0.76 (-0.19 to 1.71)	0.117
SSOB at 2021	M3	296	—	0.33 (-0.07 to 0.72)	0.105
SSOB at 2021	M4	296	—	0.33 (-0.08 to 0.74)	0.119
SSOB at 2021	M2-L	296	—	0.78 (-0.19 to 1.76)	0.114

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S8 HOMA-IR per doubling in the renal-restricted cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	296	34	0.99 (0.64 to 1.57)	0.981
DOI	M1	296	34	0.87 (0.55 to 1.41)	0.570
DOI	M2	296	34	0.74 (0.45 to 1.27)	0.273
DOI	M3	296	34	0.69 (0.40 to 1.22)	0.200
DOI	M4	296	34	0.68 (0.36 to 1.26)	0.222
DOI	M2-L	296	34	0.74 (0.42 to 1.33)	0.317
DISH	M0	296	29	1.48 (0.90 to 2.43)	0.120
DISH	M1	296	29	1.30 (0.78 to 2.23)	0.317
DISH	M2	296	29	0.88 (0.51 to 1.57)	0.656
DISH	M3	296	29	0.80 (0.46 to 1.43)	0.442
DISH	M4	296	29	0.88 (0.49 to 1.60)	0.681
DISH	M2-L	296	29	0.91 (0.51 to 1.64)	0.744
SSOB at 2021	M0	296	—	0.34 (-0.15 to 0.83)	0.169
SSOB at 2021	M1	296	—	0.14 (-0.26 to 0.53)	0.494
SSOB at 2021	M2	296	—	0.02 (-0.43 to 0.46)	0.947
SSOB at 2021	M3	296	—	-0.15 (-0.36 to 0.07)	0.174
SSOB at 2021	M4	296	—	-0.15 (-0.37 to 0.08)	0.192
SSOB at 2021	M2-L	296	—	-0.02 (-0.46 to 0.42)	0.926

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S9 TyG ≥8.67 in the renal-restricted cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	296	34	2.94 (1.31 to 6.36)	0.010
DOI	M1	296	34	2.02 (0.84 to 4.65)	0.114
DOI	M2	296	34	1.89 (0.78 to 4.46)	0.157
DOI	M3	296	34	1.54 (0.50 to 4.38)	0.439
DOI	M4	296	34	1.58 (0.52 to 4.54)	0.410
DOI	M2-L	296	34	1.94 (0.80 to 4.60)	0.141
DISH	M0	296	29	1.82 (0.70 to 4.27)	0.205
DISH	M1	296	29	1.33 (0.49 to 3.34)	0.557
DISH	M2	296	29	0.92 (0.32 to 2.42)	0.875
DISH	M3	296	29	0.66 (0.19 to 1.93)	0.462
DISH	M4	296	29	0.67 (0.19 to 1.95)	0.477
DISH	M2-L	296	29	0.89 (0.30 to 2.34)	0.814
SSOB at 2021	M0	296	—	1.63 (0.02 to 3.25)	0.047
SSOB at 2021	M1	296	—	0.77 (-0.67 to 2.21)	0.295
SSOB at 2021	M2	296	—	0.66 (-0.78 to 2.11)	0.366
SSOB at 2021	M3	296	—	0.15 (-0.46 to 0.76)	0.629
SSOB at 2021	M4	296	—	0.15 (-0.47 to 0.78)	0.624
SSOB at 2021	M2-L	296	—	0.65 (-0.81 to 2.11)	0.381

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S10 TyG ≥P75 in the renal-restricted cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	296	34	3.26 (1.57 to 6.75)	0.002
DOI	M1	296	34	2.07 (0.94 to 4.53)	0.071
DOI	M2	296	34	1.96 (0.85 to 4.53)	0.115
DOI	M3	296	34	1.54 (0.58 to 4.01)	0.379
DOI	M4	296	34	1.52 (0.55 to 4.09)	0.412
DOI	M2-L	296	34	2.06 (0.89 to 4.81)	0.092
DISH	M0	296	29	2.87 (1.30 to 6.22)	0.009
DISH	M1	296	29	2.06 (0.89 to 4.74)	0.092
DISH	M2	296	29	1.39 (0.56 to 3.36)	0.467
DISH	M3	296	29	1.15 (0.43 to 2.95)	0.771
DISH	M4	296	29	1.09 (0.41 to 2.80)	0.855
DISH	M2-L	296	29	1.35 (0.55 to 3.26)	0.509
SSOB at 2021	M0	296	—	1.84 (0.59 to 3.09)	0.004
SSOB at 2021	M1	296	—	0.88 (-0.27 to 2.02)	0.133
SSOB at 2021	M2	296	—	0.79 (-0.39 to 1.97)	0.187
SSOB at 2021	M3	296	—	0.49 (0.01 to 0.97)	0.046
SSOB at 2021	M4	296	—	0.49 (-0.01 to 0.99)	0.053
SSOB at 2021	M2-L	296	—	0.85 (-0.37 to 2.06)	0.171

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S11 TyG per SD in the renal-restricted cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	296	34	1.49 (1.07 to 2.06)	0.018
DOI	M1	296	34	1.09 (0.72 to 1.62)	0.666
DOI	M2	296	34	1.03 (0.66 to 1.57)	0.892
DOI	M3	296	34	0.76 (0.42 to 1.31)	0.340
DOI	M4	296	34	0.79 (0.43 to 1.37)	0.421
DOI	M2-L	296	34	1.06 (0.68 to 1.61)	0.807
DISH	M0	296	29	1.57 (1.12 to 2.22)	0.010
DISH	M1	296	29	1.36 (0.88 to 2.08)	0.159
DISH	M2	296	29	1.11 (0.69 to 1.75)	0.654
DISH	M3	296	29	0.96 (0.56 to 1.59)	0.884
DISH	M4	296	29	0.97 (0.56 to 1.58)	0.895
DISH	M2-L	296	29	1.11 (0.68 to 1.74)	0.665
SSOB at 2021	M0	296	—	0.82 (0.32 to 1.31)	0.001
SSOB at 2021	M1	296	—	0.16 (-0.26 to 0.58)	0.452
SSOB at 2021	M2	296	—	0.08 (-0.35 to 0.52)	0.711
SSOB at 2021	M3	296	—	-0.07 (-0.24 to 0.11)	0.457
SSOB at 2021	M4	296	—	-0.08 (-0.26 to 0.10)	0.399
SSOB at 2021	M2-L	296	—	0.10 (-0.35 to 0.56)	0.655

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S12 HOMA-IR ≥P75 in the renal-restricted cohort

Outcome	Model	N	Events	Estimate (95% CI)	Nominal P
DOI	M0	296	34	1.57 (0.71 to 3.30)	0.255
DOI	M1	296	34	1.14 (0.48 to 2.56)	0.761
DOI	M2	296	34	0.97 (0.38 to 2.37)	0.946
DOI	M3	296	34	0.70 (0.22 to 2.02)	0.517
DOI	M4	296	34	0.82 (0.25 to 2.45)	0.723
DOI	M2-L	296	34	1.03 (0.39 to 2.58)	0.955
DISH	M0	296	29	2.81 (1.28 to 6.08)	0.011
DISH	M1	296	29	1.99 (0.86 to 4.55)	0.109
DISH	M2	296	29	1.23 (0.48 to 3.05)	0.660
DISH	M3	296	29	1.00 (0.37 to 2.60)	0.998
DISH	M4	296	29	1.14 (0.40 to 3.10)	0.800
DISH	M2-L	296	29	1.29 (0.48 to 3.35)	0.603
SSOB at 2021	M0	296	—	1.22 (-0.07 to 2.51)	0.063
SSOB at 2021	M1	296	—	0.63 (-0.49 to 1.76)	0.267
SSOB at 2021	M2	296	—	0.52 (-0.76 to 1.80)	0.427
SSOB at 2021	M3	296	—	-0.19 (-0.69 to 0.30)	0.445
SSOB at 2021	M4	296	—	-0.18 (-0.69 to 0.32)	0.471
SSOB at 2021	M2-L	296	—	0.44 (-0.78 to 1.66)	0.477

DOI and DISH estimates are odds ratios; SSOB21 estimates are mean follow-up score differences. Binary references are below the stated cut point. P75 cut points are TyG 8.45 and HOMA-IR 1.08 (rounded for display); unrounded full-cohort values were used in both populations. M0, unadjusted; M1, age and sex; M2, M1 plus waist circumference; M3, M2 plus baseline SSOB; M4, M3 plus smoking, alcohol, and exercise; M2-L, M2 plus the same lifestyle variables without baseline SSOB. Exposures were fitted separately.

Table S13 Inter-reader agreement for SSOB and six-year change

Assessment	Region	N	Reader A mean	Reader B mean	Mean B − A	ICC(A,1) (95% CI)	Exact n (%)	Within 1 point n (%)
2015	Total	307	1.257	1.267	0.010	0.997 (0.991–0.999)	296 (96.4)	305 (99.3)
2015	Cervical	307	0.827	0.827	0.000	0.992 (0.980–0.998)	299 (97.4)	305 (99.3)
2015	Lumbar	307	0.430	0.440	0.010	0.998 (0.992–1.000)	304 (99.0)	307 (100.0)
2021	Total	307	1.827	1.684	-0.143	0.985 (0.965–0.993)	259 (84.4)	292 (95.1)
2021	Cervical	307	1.114	0.997	-0.117	0.962 (0.916–0.983)	271 (88.3)	294 (95.8)
2021	Lumbar	307	0.713	0.687	-0.026	0.992 (0.981–0.997)	293 (95.4)	305 (99.3)
2015–2021 change	Total	307	0.570	0.417	-0.153	0.914 (0.844–0.953)	252 (82.1)	292 (95.1)
2015–2021 change	Cervical	307	0.287	0.169	-0.117	0.856 (0.727–0.928)	264 (86.0)	291 (94.8)
2015–2021 change	Lumbar	307	0.283	0.248	-0.036	0.956 (0.923–0.978)	290 (94.5)	305 (99.3)

SSOB is the sum of 13 segmental scores (0–3 each): seven cervical and six lumbar segments. Change = 2021 minus 2015. ICC(A,1): two-way mixed-effects, single-measurement, absolute-agreement intraclass correlation coefficient; the two readers were fixed and participants random. Confidence intervals are percentile intervals from 5,000 participant-level bootstrap resamples (seed 20260927). Agreement compares the original readings before adjudication. CI, confidence interval; SSOB, spinal segmental ossification burden.

Table S14 Inter-reader agreement for segmental ossification scores

2015 | 307 participants

Region or segment	Paired scores	Exact agreement n (%)	Linear weighted κ (95% CI)	Quadratic weighted κ (95% CI)	Valid bootstrap replicates[1]
All 13 segments	3991	3980 (99.7)	0.982 (0.969–0.992)	0.985 (0.971–0.995)	5000
Cervical	2149	2141 (99.6)	0.979 (0.961–0.992)	0.980 (0.957–0.994)	5000
Lumbar	1842	1839 (99.8)	0.988 (0.971–1.000)	0.994 (0.982–1.000)	5000
C1/2	307	305 (99.3)	0.498 (0.000–1.000)	0.498 (0.000–1.000)	4741
C2/3	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	4968
C3/4	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	5000
C4/5	307	304 (99.0)	0.946 (0.872–1.000)	0.939 (0.846–1.000)	5000
C5/6	307	305 (99.3)	0.984 (0.961–1.000)	0.988 (0.969–1.000)	5000
C6/7	307	306 (99.7)	0.992 (0.970–1.000)	0.994 (0.977–1.000)	5000
C7/T1	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	5000
T12/L1	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	5000
L1/2	307	305 (99.3)	0.958 (0.879–1.000)	0.971 (0.906–1.000)	5000
L2/3	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	5000
L3/4	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	5000
L4/5	307	306 (99.7)	0.969 (0.864–1.000)	0.986 (0.928–1.000)	5000
L5/S1	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	4999

2021 | 307 participants

Region or segment	Paired scores	Exact agreement n (%)	Linear weighted κ (95% CI)	Quadratic weighted κ (95% CI)	Valid bootstrap replicates[1]
All 13 segments	3991	3921 (98.2)	0.918 (0.889–0.941)	0.938 (0.905–0.961)	5000
Cervical	2149	2096 (97.5)	0.892 (0.849–0.927)	0.908 (0.854–0.946)	5000
Lumbar	1842	1825 (99.1)	0.956 (0.930–0.976)	0.974 (0.954–0.987)	5000
C1/2	307	307 (100.0)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	4791
C2/3	307	305 (99.3)	0.955 (0.872–1.000)	0.978 (0.932–1.000)	5000
C3/4	307	303 (98.7)	0.914 (0.781–0.982)	0.958 (0.871–0.992)	5000
C4/5	307	295 (96.1)	0.850 (0.752–0.930)	0.846 (0.706–0.945)	5000
C5/6	307	289 (94.1)	0.874 (0.810–0.928)	0.898 (0.829–0.948)	5000
C6/7	307	300 (97.7)	0.944 (0.898–0.979)	0.963 (0.929–0.987)	5000
C7/T1	307	297 (96.7)	0.782 (0.607–0.906)	0.808 (0.611–0.930)	5000
T12/L1	307	303 (98.7)	0.962 (0.917–0.992)	0.981 (0.957–0.996)	5000
L1/2	307	305 (99.3)	0.963 (0.894–1.000)	0.964 (0.883–1.000)	5000
L2/3	307	303 (98.7)	0.950 (0.896–0.989)	0.975 (0.947–0.995)	5000
L3/4	307	305 (99.3)	0.962 (0.892–1.000)	0.978 (0.930–1.000)	5000
L4/5	307	304 (99.0)	0.928 (0.823–1.000)	0.966 (0.903–1.000)	5000

Region or segment	Paired scores	Exact agreement n (%)	Linear weighted κ (95% CI)	Quadratic weighted κ (95% CI)	Valid bootstrap replicates[1]
L5/S1	307	305 (99.3)	0.953 (0.877–1.000)	0.974 (0.920–1.000)	5000

Primary weighting was linear: $1 - |i - j|/3$; quadratic weighting, $1 - (i - j)^2/9$, is supplementary ($i,j = 0,1,2,3$). Pooled estimates retain all segments within each participant during resampling. Percentile 95% CIs use 5,000 participant-level bootstrap resamples. [1]Valid replicate counts refer to linear weighted κ; replicates with no rating variation were omitted. A bootstrap interval of 1.000–1.000 in a sample with no observed disagreements does not establish perfect population agreement. Thoracic segments available only in 2021 were excluded.

Table S15 Inter-reader agreement for ossification progression classifications

A Three-category classification

Reader A / Reader B	NOI	EOI	DOI	Total
NOI	201	4	2	207
EOI	8	52	3	63
DOI	1	4	32	37
Total	210	60	37	307

Exact agreement: 285/307 (92.8%); unweighted Cohen's κ = 0.852 (0.789–0.907).

B Binary DOI classification

Reader A / Reader B	non-DOI	DOI	Total
non-DOI	265	5	270
DOI	5	32	37
Total	270	37	307

Exact agreement: 297/307 (96.7%); unweighted Cohen's κ = 0.846 (0.742–0.929).

Classification was derived independently from each reader's 2015 and 2021 scores using the study definitions, with DOI taking precedence over EOI. Both readers identified 37 DOI cases; 32 were common to both, with five classified as DOI only by each reader. Percentile 95% CIs use 5,000 participant-level bootstrap resamples. NOI, no ossification increase; EOI, equivocal ossification increase; DOI, definite ossification increase. Diagnostic concordance reported by the investigators was not used to infer a numerical κ for diagnostic categories.

Supplementary Figures

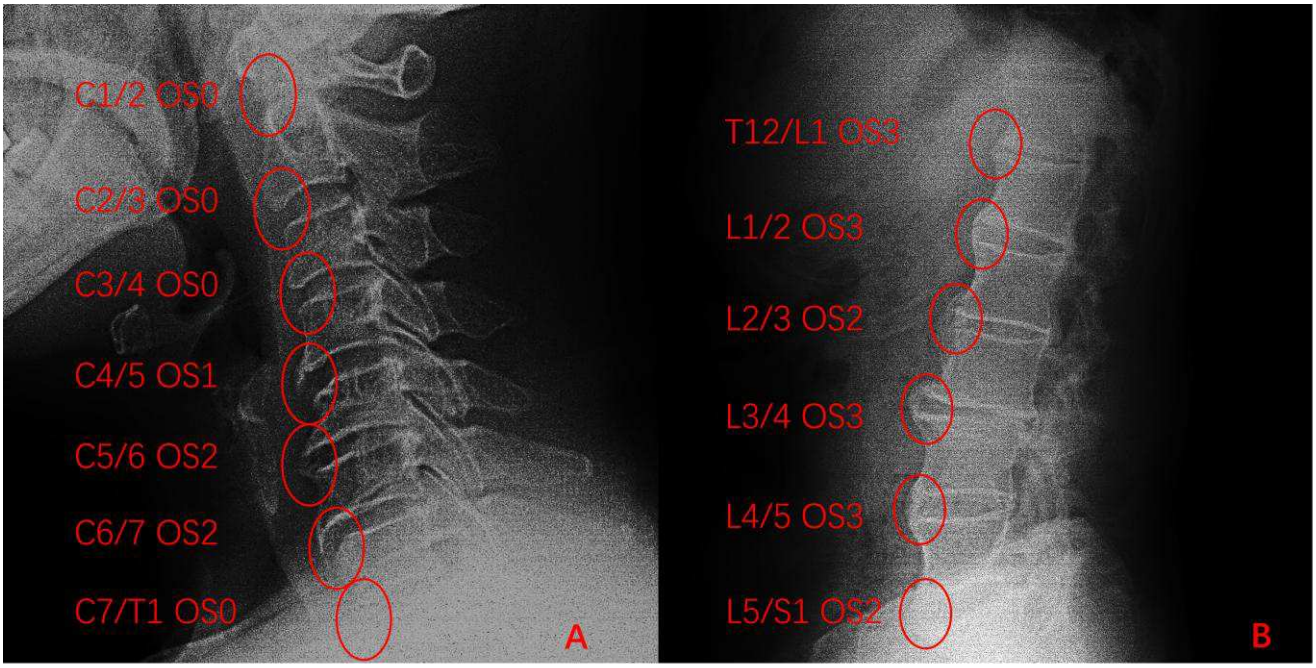


Fig. S1 Radiographic examples for SSOB calculating method.

(A) demonstrates a lateral radiograph of the cervical spine. No significant anterior vertebral ossification was observed at segments C1-4. The ossification grade was OS1 at C4/5, OS2 at C5/6, OS2 at C6/7, and OS0 at C7/T1. The sum of the ossification grades across all cervical segments constitutes the cervical SSOB, with a total value of 5. (B) illustrates a lateral radiograph of the lumbar spine. The ossification grades were 3 at T12/L1, L1/2, L3/4, and L4/5, and 2 at L2/3 and L5/S1. The lumbar SSOB was calculated as the sum of the grades for all lumbar segments, yielding a value of 16.

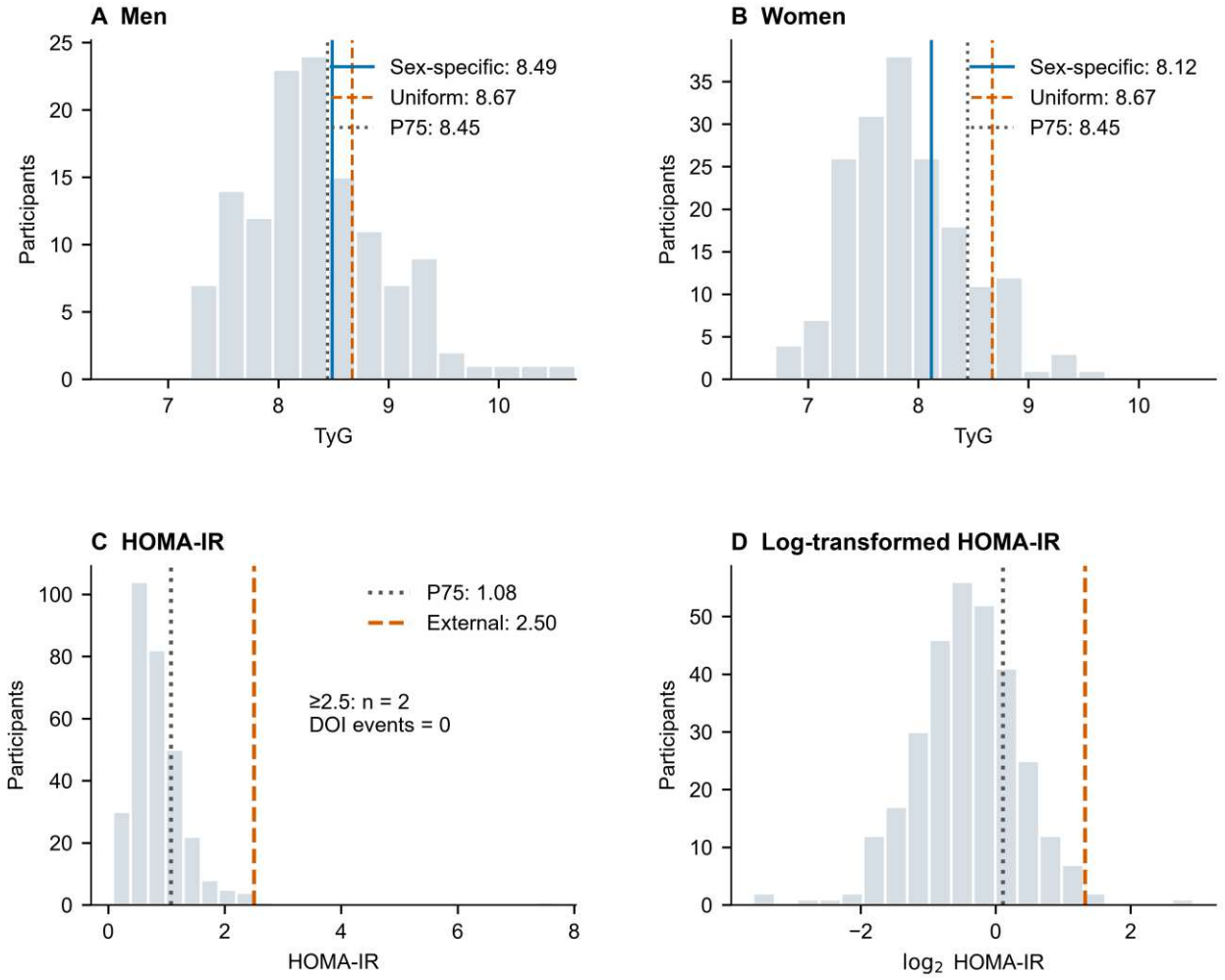
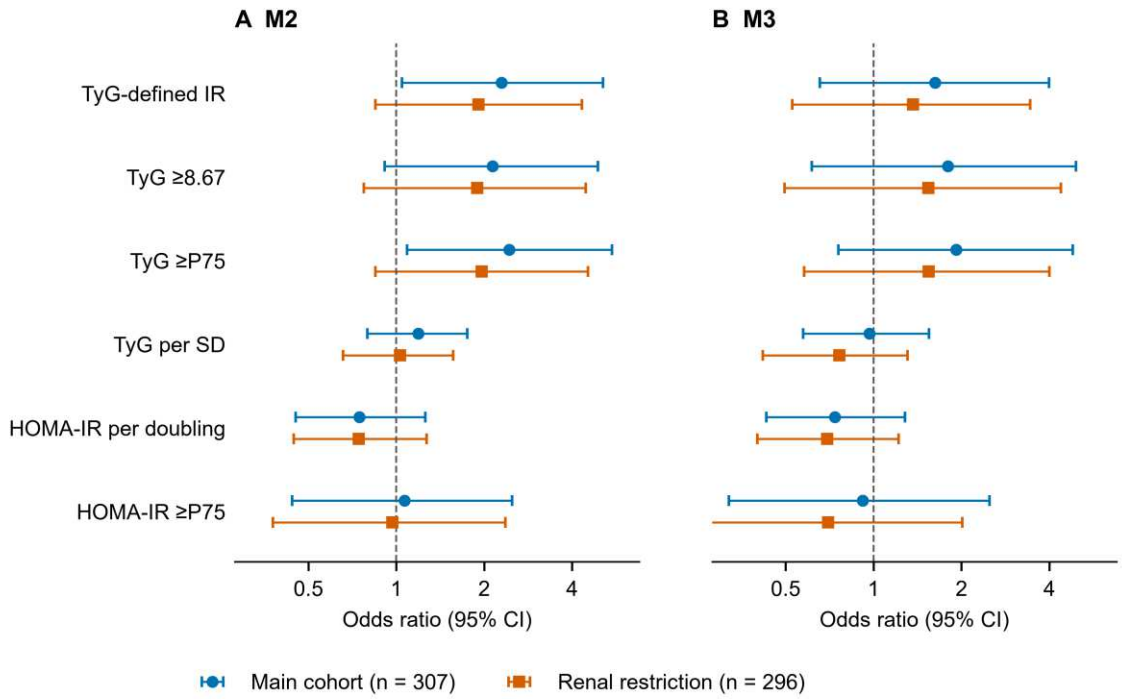


Fig. S2 Baseline exposure distributions and thresholds. (A) TyG in men. (B) TyG in women. Sex-specific thresholds are shown alongside the uniform threshold of 8.67 and cohort P75 of 8.45. (C) HOMA-IR on its original scale. (D) HOMA-IR on the $\log_2$ scale. The HOMA-IR P75 is 1.08; two participants meet the external threshold of 2.5. Quantiles are rounded for display



DOI events: 37 in the main cohort and 34 after renal restriction.
Exposure units differ; each row represents a separate model.

Fig. S3 DOI associations across exposure definitions in the full and renal-restricted cohorts. (A) M2 adjusts for age, sex, and waist circumference. (B) M3 additionally adjusts for baseline SSOB. Points and lines represent odds ratios and 95% confidence intervals; binary and continuous contrasts are modeled separately. The full cohort includes 307 participants and 37 DOI events; the renal-restricted cohort includes 296 participants and 34 events

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMethods.docx