Journal of Experimental and Integrative Medicine
www.jeim.org

Original Article



Pancreatic volume loss on routine restaging CT as an early-warning biomarker of checkpoint inhibitor–associated autoimmune diabetes: Protocol for a multicentre nested case–control study (PANC-WATCH)

Christopher Liddle[1], Nicole Fewings[2], Matteo Carlino[2,3], David A. Brown[1,2], Richard A. Scolyer[1], Roderick Clifton-Bligh[1]

[1]Westmead Institute for Medical Research, University of Sydney, NSW, Sydney, Australia.

[2]Centre for Diabetes, Obesity and Endocrinology Research, Westmead Institute for Medical Research, Westmead, NSW, Sydney, Australia.

[3]Department of Diabetes and Endocrinology, Westmead Hospital, NSW, Sydney, Australia.

Address for correspondence:
Richard A. Scolyer,
Westmead Institute for
Medical Research,
University of Sydney,
NSW, Sydney, Australia.
E-mail:
richard.scolyer@health.nsw.gov.au

Received: August 29, 2025
Accepted: June 27, 2026
Published: July 03, 2026

ABSTRACT

Checkpoint inhibitor–associated autoimmune diabetes (CIADM) affects about 1% of patients treated with immune checkpoint inhibitors and presents with diabetic ketoacidosis in about two-thirds. The pancreas is smaller before treatment and markedly smaller at diagnosis in CIADM, but it is unknown when shrinkage begins, because few patients have been studied with scans taken during treatment. To determine whether pancreatic volume measured automatically on routine restaging CT falls before CIADM is diagnosed, and whether a simple threshold can provide early warning. Retrospective multicentre nested case–control study. Sixty CIADM cases will be matched to 240 ICI-treated controls by incidence-density sampling on centre, regimen, tumour group, age, sex and time since ICI start. Every CT from baseline to the index date will be segmented with an open-source deep-learning model and checked by a blinded radiologist. Pancreatic volume will be indexed to body surface area. Linear and piecewise mixed-effects models will compare volume trajectories and estimate when decline accelerates. Thresholds of 10%, 15% and 20% decline will be evaluated with conditional logistic regression, time-dependent ROC curves, lead time and decision curves. Automated and manual volumetry will be compared in 20% of scans. Sixty cases provide 84% power to detect a 6-point difference in percentage change. A positive result would allow radiologists to flag impending CIADM on scans that patients already receive, creating a window to prevent ketoacidosis.

KEY WORDS: Immune checkpoint inhibitors, autoimmune diabetes, diabetic ketoacidosis, pancreatic volume, computed tomography, deep learning segmentation, nested case–control study

Introduction

Immune checkpoint inhibitors (ICIs) that block programmed cell death 1 (PD-1), its ligand PD-L1, or cytotoxic T-lymphocyte antigen 4 (CTLA-4) have changed the outlook for many advanced cancers, but they also cause immune-related adverse events that can affect almost any organ [1, 2]. Checkpoint inhibitor–associated autoimmune diabetes mellitus (CIADM) is one of the least common of these events, but one of the most serious. In a claims database of more than 30,000 ICI-treated patients, insulin-requiring diabetes developed in 0.86%, and the risk was higher with combined anti-CTLA-4 and anti-PD-1 therapy [3]. Single-centre series report similar rates of about 1% to 2% [4, 5].

CIADM differs from classical type 1 diabetes. It usually starts abruptly, often without islet autoantibodies, and causes near-complete and permanent loss of insulin secretion [5, 6]. A systematic review of 192 patients found that 69.7% presented with diabetic ketoacidosis (DKA), and only 40.4% had type 1 diabetes autoantibodies [7]. Because the onset is sudden and frequently unrecognised until ketoacidosis develops, the clinical priority is not only to predict who will develop CIADM, but to detect it *before* it becomes an emergency. Current guidelines recommend checking glucose during ICI therapy, but they offer no tool that identifies impending beta-cell failure [2].

The pancreas shrinks in CIADM

Several independent case series have shown that the pancreas becomes smaller in CIADM (Table 1). Byun and colleagues described marked radiological pancreatic atrophy in 18 patients [8]. In a multicentre Australian series, pancreatic volume was already lower before ICI treatment in patients who later developed CIADM and fell by a mean of 41% from pre-treatment to six months after diagnosis [9]. In a Taiwanese cohort of 2829 ICI-treated patients, the ten who developed diabetes showed a median fall of 19.9% from baseline at diagnosis and 37.1% at last follow-up [10]. Most recently, a nested case–control study in metastatic melanoma reported that baseline pancreatic volume was 60 versus 73 mL in future cases and controls, and that pancreatic volume combined with low-level islet autoantibodies predicted CIADM [11]. A similar picture is seen in classical type 1 diabetes, in which pancreatic volume is reduced before diagnosis and continues to decline after it [12, 13].

The unanswered question

None of these studies could show *when* the pancreas begins to shrink. They compared volume before treatment with volume at or after diagnosis, and the most recent study noted that most patients had no scan on treatment before onset [11]. Yet patients receiving ICIs have restaging CT every 8 to 12 weeks as part of routine cancer care. If pancreatic volume falls

Table 1: Published evidence on pancreatic volume in checkpoint inhibitor–associated diabetes

Study	Population	Imaging	Main finding on pancreatic volume
Byun, 2020 [8]	18 ICI-DM cases, mixed tumours, USA	Routine CT, qualitative	Significant pancreatic atrophy after diagnosis
Wu, 2024 [9]	17 CIADM cases with volumetry and ICI-treated comparison groups, Australia	CT volumetry	Lower volume before ICI; mean 41% fall from pre-treatment to 6 months after diagnosis
Wei, 2024 [10]	10 ICI-DM cases among 2829 ICI-treated patients, Taiwan	Manual CT segmentation	Median fall of 19.9% at diagnosis and 37.1% at last follow-up
Wu, 2026 [11]	14 cases, 28 matched controls, melanoma, Australia	CT volumetry, single reader	Baseline volume 60 vs 73 mL; large fall at diagnosis; few on-treatment scans before onset

measurably on these scans weeks before hyperglycaemia, the information needed to prevent DKA is already being collected and simply not being measured. Automated deep-learning segmentation now allows the pancreas to be measured on routine CT without manual tracing [14, 15, 16], which makes a large multicentre analysis feasible.

Study questions

PANC-WATCH (*PANCreatic volume WATCH during checkpoint inhibitor therapy*) will answer four questions:

1. Does pancreatic volume decline faster in patients who develop CIADM than in matched ICI-treated controls during the months *before* diagnosis?
2. Is there a point of acceleration, and how long before diagnosis does it occur?
3. Can a simple rule (for example, a fall of 15% or more from baseline) flag future CIADM with useful sensitivity, specificity and lead time?
4. Is automated volumetry on routine CT reliable enough for clinical use?

Materials and Methods

The study is organised into four work packages (WPs), shown in Figure 6. It will be reported according to STROBE for observational studies [17] and TRIPOD+AI for the early-warning rule [18].

WP1. Study population and case–control sampling

Design and setting. PANC-WATCH is a retrospective nested case–control study within the ICI-treated populations of participating cancer centres in Pakistan and partner centres abroad (to be named). The source cohort includes all adults who started anti-PD-1 or anti-PD-L1 therapy, with or without anti-CTLA-4, between January 2016 and December 2025, and who had a CT of the abdomen within 90 days before the first dose and at least one CT during treatment.

Cases. Cases meet the CIADM definition used in previous studies [7, 9]: (i) previous ICI therapy; (ii) new hyperglycaemia, with blood glucose $\geq$ 11.1 mmol/L or HbA1c $\geq$ 6.5%; and (iii) insulin deficiency, with DKA or C-peptide $<$ 0.4 nmol/L within one month. Patients with insulin-treated diabetes before ICI are excluded. Pre-existing type 2 diabetes not treated with insulin is allowed and recorded, because it is a known risk factor [3]. The *index date* is the date of CIADM diagnosis.

Table 2: Matching factors for controls (4 per case)

Factor	Matching rule
Centre	Same centre (controls scanning protocol)
ICI regimen	Anti-PD-1/PD-L1 alone vs combined with anti-CTLA-4
Tumour group	Melanoma; lung; kidney/urothelial; other
Age	Within 5 years
Sex	Same
Time since ICI start	At risk at the case's index time (risk set)
Calendar period	First ICI dose within 2 years of the case

Controls. Controls are selected by incidence-density (risk-set) sampling (Figure 1). For each case, four controls are drawn from patients in the same source cohort who were still on follow-up, alive and free of diabetes at the same time since ICI start as the case at its index date. This time becomes the control's *pseudo-index date*. Controls are matched on the factors listed in Table 2. Risk-set sampling ensures that cases and controls have had the same exposure time to ICIs and the same opportunity to have restaging scans, and makes the odds ratio an unbiased estimate of the hazard ratio.

Exclusions. Patients are excluded if they had pancreatic primary cancer or metastasis, pancreatic surgery, acute pancreatitis before ICI, or no baseline CT that includes the whole pancreas. Pancreatitis or lipase elevation *during* ICI treatment is not an exclusion criterion, because it may be part of the pathway to CIADM [4, 9]; it is recorded and examined in sensitivity analyses.

WP2. Image retrieval and automated volumetry

All CT scans that include the pancreas from 90 days before the first ICI dose to the index or pseudo-index date are retrieved from the picture archiving and communication system, de-identified, and processed with the same pipeline (Figure 2). Each scan must meet minimum quality criteria (Table 3).

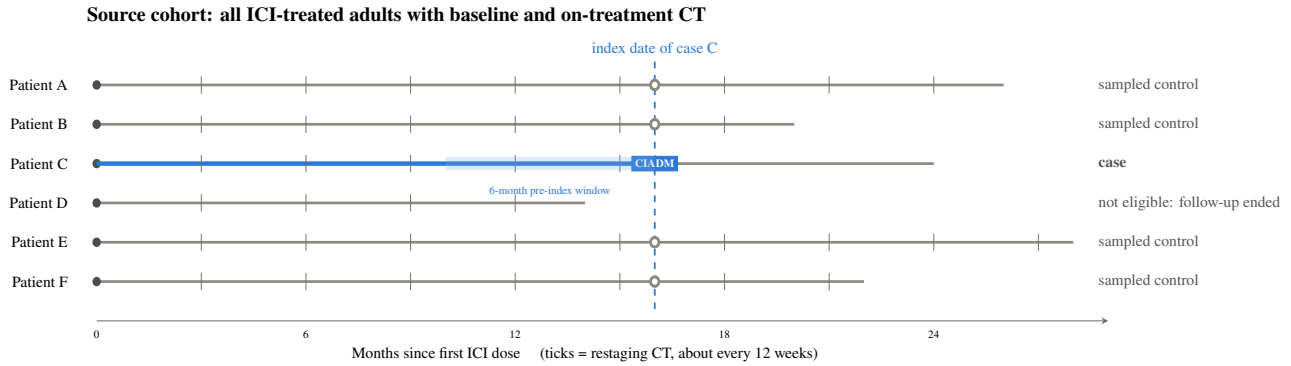


Figure 1: Nested case–control design with incidence-density sampling. When a case develops CIADM, four controls are drawn from patients still at risk at the same time since ICI start. All CT scans up to the index or pseudo-index date are measured. The shaded segment marks the six-month window before diagnosis used for the primary acceleration analysis.

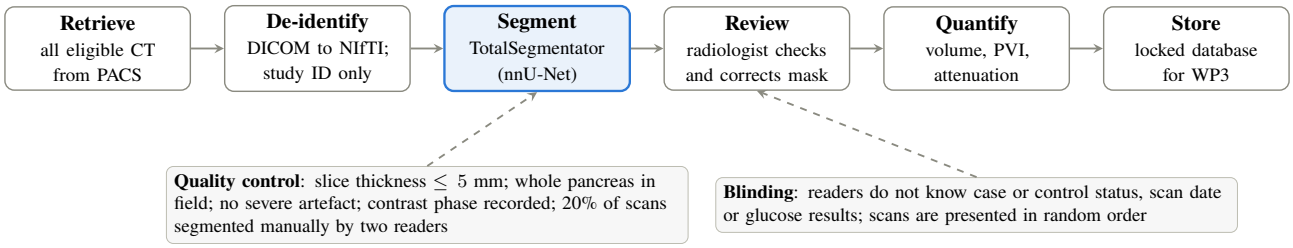


Figure 2: Imaging pipeline for automated pancreatic volumetry. PACS = picture archiving and communication system; PVI = pancreatic volume index (volume divided by body surface area).

Segmentation. The pancreas is segmented automatically with TotalSegmentator, a publicly available nnU-Net–based model trained on more than 1200 routine clinical CT examinations [14, 15]. A radiologist reviews every mask and corrects errors, particularly at the pancreatic head and duodenal border, where automated segmentation is least accurate [16]. Readers are blinded to case status, scan date and glucose results.

Measurements. For each scan we record (i) total pancreatic volume (mL); (ii) the pancreatic volume index (PVI), defined as volume divided by body surface area, to adjust for body size and weight loss [19, 13]; (iii) mean parenchymal attenuation in Hounsfield units, as a marker of oedema or fat; and (iv) the percentage change from baseline,

$$\Delta V_t\,\% = 100 \times \frac{V_t - V_0}{V_0}, \tag{1}$$

where V_0 is the baseline volume and V_t the volume on a scan at time t. Pancreatic volume falls slowly with age and varies with body fat [19], so analyses always compare cases with time-matched controls.

Reliability. A random 20% of scans will also be segmented manually by two radiologists. Agreement between automated and manual volumes, and between the two readers, will be measured by the intraclass correlation coefficient [20], the Dice similarity coefficient and Bland–Altman limits of agreement [21]. Automated volumetry will be considered adequate if the ICC is at least 0.90 and the limits of agreement for percentage change are within ± 10 percentage points.

WP3. Statistical analysis

The analysis plan is summarised in Table 4.

Table 3: CT acceptance and quality-control criteria

Item	Requirement
Coverage	Entire pancreas from head to tail
Slice thickness	≤ 5 mm; 1–3 mm preferred
Phase	Portal venous preferred; non-contrast and arterial accepted and recorded
Artefact	No major motion or metal artefact over pancreas
Pathology	No pancreatic mass, metastasis, or surgery
Timing	Baseline within 90 days before first ICI; on-treatment scans up to index date
Segmentation	Mask reviewed; corrected if boundary error > 2 slices

Primary analysis: volume trajectories. The primary outcome is the trajectory of PVI over time before the index date. We will fit a linear mixed-effects model [22] with time before index (months), group (case or control) and their interaction as fixed effects, adjusted for age, sex, baseline body mass index, contrast phase and centre, with random intercepts and slopes for each patient and a random intercept for each matched set. To test whether decline *accelerates* before diagnosis, a piecewise (change-point) model will allow a different slope in the final six months before the index date (Figure 3). The change point will be estimated by profile likelihood over the range 1 to 9 months, with a bootstrap confidence interval.

Secondary analysis: early-warning rule. For each patient, the *last eligible scan* is the latest CT at least 14 days before the index or pseudo-index date, so that a flag on that scan

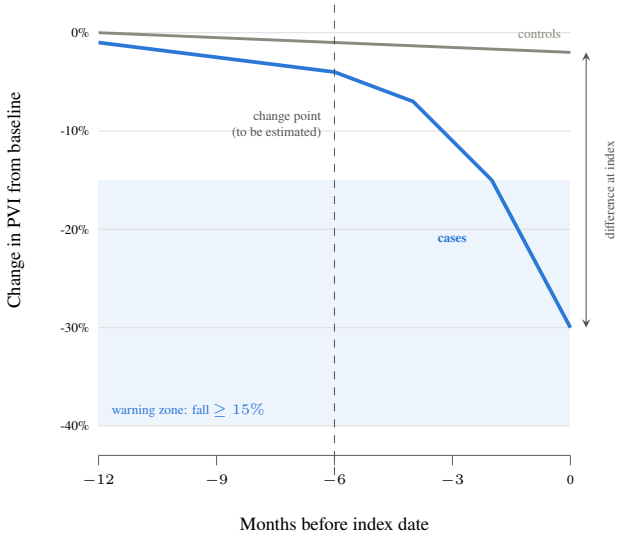


Figure 3: Hypothesised (not observed) trajectories of pancreatic volume index (PVI) before diagnosis. Cases are expected to show slow decline followed by acceleration; controls change little. The analysis will estimate the change point and the slopes.

could have led to action. The association between percentage change in PVI on that scan and CIADM will be estimated by conditional logistic regression within matched sets. Discrimination over time will be assessed with time-dependent ROC curves [23], and pre-specified thresholds of 10%, 15% and 20% decline will be reported with sensitivity, specificity, predictive values and the median *lead time* from the first flagged scan to diagnosis. Because cases are over-represented in a case–control sample, predictive values will be recalculated for a CIADM incidence of 1% [3]. Clinical usefulness will be examined by decision-curve analysis [24]. A combined model with baseline PVI, percentage change and ICI regimen will be developed with penalised (ridge) regression and internally validated by bootstrap, reporting optimism-corrected discrimination and calibration slope [18, 25].

Exploratory analyses. These include parenchymal attenuation; the baseline threshold of 75 mL suggested by Wu and colleagues [11]; the association of pre-index volume loss with DKA at presentation; and subgroups by tumour type, ICI regimen and lipase elevation.

Missing data. Patients differ in the number and timing of scans; mixed models use all available scans without imputation, assuming data are missing at random. Missing covariates will be handled by multiple imputation. A sensitivity analysis will restrict to patients with at least two on-treatment scans before the index date.

Sample size

The sample size is driven by the primary comparison of percentage change in PVI on the last eligible scan. Published series suggest a fall of about 20% by diagnosis in cases [10, 9]; before diagnosis the fall will be smaller. We therefore assumed a mean fall of 12% (SD 15) in cases and 3% (SD 10) in controls, a difference of 9 percentage points, and also considered a more conservative difference of 6 points. With

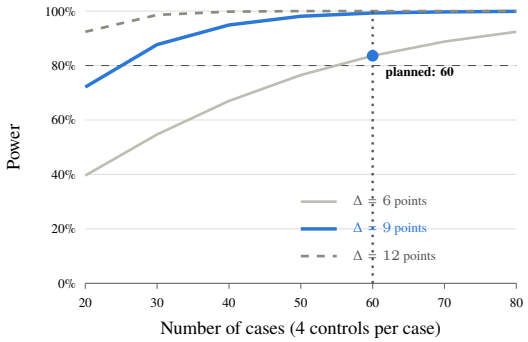


Figure 4: Power to detect a difference (Δ) in mean percentage change of pancreatic volume index between cases and controls (SD 15 in cases, 10 in controls; two-sided $\alpha = 0.05$).

four controls per case and a two-sided α of 0.05, the number of cases needed is

$$n_{\text{cases}} = \frac{\left(z_{1-\alpha 2} \;\; z_{1-\beta}\right)^2 \left(\sigma_1^2 \;\; \sigma_0^2 k\right)}{\Delta^2}, \qquad (2)$$

where $k = 4$ is the number of controls per case. For $\Delta = 9$, 33 cases give 90% power; for $\Delta = 6$, 55 cases give 80% power. We will therefore aim for **60 cases and 240 controls**, which gives 84% power for the conservative difference and over 99% for the expected one (Figure 4, Table 5). With 60 cases, a combined model with three predictors has 20 events per predictor, consistent with current guidance on minimum sample size for prediction models [25].

At an incidence of about 0.9% [3], 60 cases require a source population of roughly 6700 ICI-treated patients. This is why the study is planned as a consortium of at least four to six centres.

WP4. From measurement to clinical action

If the early-warning rule performs well, it will be translated into a simple pathway for radiologists and oncologists (Figure 5). A flagged scan would prompt same-week capillary glucose, blood ketones, HbA1c and C-peptide, patient education on symptoms of hyperglycaemia, and review by an endocrinologist. This WP will draft the pathway and the wording of a structured radiology report, to be tested prospectively in a later study.

Timeline

The project will take 24 months (Figure 6). Image processing can start at each centre as soon as its cases and controls are identified.

Ethics and data governance

The study uses existing clinical records and images, so a waiver of individual consent will be requested from each institutional review board. Images will be de-identified at each centre; only study identifiers and derived measurements will be shared. The segmentation software is open source and runs locally, so no images leave the hospital network. The study follows the Declaration of Helsinki.

Expected Results

No data have been collected. The results will be presented as follows: a flow diagram of patients and scans; baseline characteristics of cases and controls (Table 6); the observed

Table 4: Statistical analysis plan

Question	Outcome or exposure	Method	Key output
Q1. Faster decline before diagnosis?	PVI over time	Linear mixed model, time × group	Difference in slope (%/month), 95% CI
Q2. When does decline accelerate?	PVI over time	Piecewise mixed model; profile likelihood	Change point (months before index), bootstrap CI
Q3. Early-warning rule	% change on last scan ≥ 14 days before index	Conditional logistic regression; time-dependent ROC; decision curve	OR per 5% fall; AUC; sensitivity, specificity, lead time; net benefit
Q3b. Combined model	Baseline PVI, % change, regimen	Ridge regression; bootstrap validation	Optimism-corrected AUC; calibration slope
Q4. Reliability	Automated vs manual; reader 1 vs 2	ICC; Dice; Bland–Altman	Agreement and limits of agreement

Table 5: Cases required (4 controls per case)

Difference Δ	80% power	90% power
6 points	55	73
9 points	25	33
12 points	14	19

SD 15 (cases) and 10 (controls); two-sided $\alpha = 0.05$.

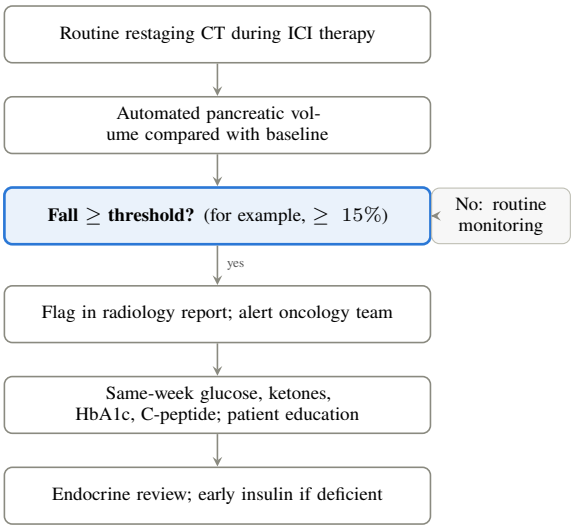


Figure 5: Proposed early-warning pathway to be developed in WP4. The threshold will be chosen from the WP3 results; the pathway itself will require prospective testing.

PVI trajectories with the estimated change point; the performance of each threshold with lead time; and reliability statistics. A positive study would show (i) a steeper fall in PVI in cases than controls before diagnosis, (ii) a measurable change point, and (iii) a threshold that flags a meaningful proportion of cases at least two weeks before diagnosis while flagging few controls.

Discussion

PANC-WATCH addresses the question left open by recent biomarker studies: not whether the pancreas is smaller in CIADM, which is now established [8, 9, 10, 11], but whether its shrinkage can be seen early enough on routine scans to change care. The approach has three practical advantages. It uses imaging that patients already receive, so it adds no cost, radiation or burden. It uses open-source automated segmentation [14], which removes the need for time-consuming manual tracing. And because the rule depends on change within the same patient, it may be less affected by differences in body size and scanner than a single baseline threshold [19].

Table 6: Result shell: characteristics of cases and controls (to be completed)

Characteristic	Cases	Controls
n	–	–
Age, years, mean (SD)	–	–
Male, n (%)	–	–
Tumour: melanoma / lung / other	–	–
Combined anti-CTLA-4, n (%)	–	–
Pre-existing type 2 diabetes, n (%)	–	–
Scans per patient, median (IQR)	–	–
Baseline PVI, mL/m^2	–	–
Change in PVI on last scan, %	–	–
Flagged (≥ 15% fall), n (%)	–	–
Lead time, days, median (IQR)	–	n/a
DKA at diagnosis, n (%)	–	n/a

Threats to validity

Table 7 lists the main threats and how the design addresses them. The most important is confounding by cancer cachexia and treatment-related weight loss, which may reduce pancreatic volume regardless of diabetes. Matching on time since ICI, normalising volume to body surface area, and adjusting for body mass index reduce this risk, and parenchymal attenuation will help to separate atrophy from fat loss.

Limitations

As a retrospective study, PANC-WATCH depends on the timing of routine scans, which differs between patients and centres. Glucose was not measured systematically in controls, so undiagnosed mild hyperglycaemia cannot be excluded. Predictive values from a case–control sample cannot be used directly in practice and must be recalculated for the true incidence. Finally, a warning rule developed here will need external validation and a prospective study to show that acting on a flag actually reduces DKA.

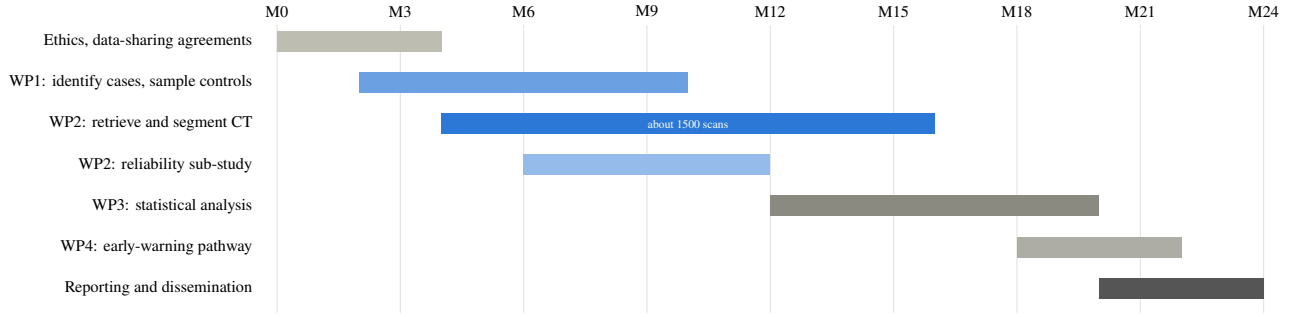


Figure 6: Gantt chart of the four work packages over 24 months. M = month.

Table 7: Threats to validity and mitigation

Threat	Why it matters	Mitigation
Too few cases	CIADM is rare (about 1%)	Multicentre consortium; power calculated for a conservative effect
Cancer cachexia	Weight loss may shrink the pancreas	Volume indexed to BSA; BMI adjustment; time-matched controls; attenuation analysis
Unequal scan schedules	Cases may be scanned more often near diagnosis	Only scans $\geq$ 14 days before index; risk-set sampling on time since ICI
Pancreatitis during ICI	Swelling can mask atrophy	Lipase and pancreatitis recorded; sensitivity analysis excluding them
Segmentation error	Pancreas borders are hard to define	Radiologist review of every mask; reliability sub-study
Contrast phase and scanner	Phase changes apparent volume	Phase recorded and adjusted; matching on centre
Reader bias	Knowing diagnosis may influence tracing	Readers blinded to status and dates; random scan order
Overfitting of warning rule	Few events	Few pre-specified predictors; penalised regression; bootstrap validation

Conclusion

Patients receiving immune checkpoint inhibitors already have regular CT scans that include the pancreas. PANC-WATCH will determine whether automated measurement of pancreatic volume on these scans reveals checkpoint inhibitor–associated diabetes before it presents as ketoacidosis. If it does, a low-cost addition to routine radiology reporting could give clinicians a window of several weeks to test, educate and treat.

Declarations

- **Registration:** [Open Science Framework or ClinicalTrials.gov; number pending].
- **Ethics approval:** [Institutional review boards of participating centres; waiver of consent requested].
- **Funding:** [to be completed].
- **Conflicts of interest:** [to be completed].
- **Author contributions:** [to be completed using the CRediT taxonomy].
- **Data availability:** Derived volumetric data and analysis code will be shared on reasonable request; images remain at each centre.

References

[1] Postow, M. A., Sidlow, R., & Hellmann, M. D. (2018). Immune-related adverse events associated with immune checkpoint blockade. *New England Journal of Medicine, 378*(2), 158–168.

[2] Haanen, J., Obeid, M., Spain, L., Carbonnel, F., Wang, Y., Robert, C., Lyon, A. R., Wick, W., Kostine, M., Peters, S., Jordan, K., Larkin, J., & ESMO Guidelines Committee. (2022). Management of toxicities from immunotherapy: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology, 33*(12), 1217–1238.

[3] Chen, X., Affinati, A. H., Lee, Y., Turcu, A. F., Henry, N. L., Schiopu, E., Qin, A., Othus, M., Clauw, D., Ramnath, N., & Zhao, L. (2022). Immune checkpoint inhibitors and risk of type 1 diabetes. *Diabetes Care, 45*(5), 1170–1176.

[4] Stamatouli, A. M., Quandt, Z., Perdigoto, A. L., Clark, P. L., Kluger, H., Weiss, S. A., Gettinger, S., Sznol, M., Young, A., Rushakoff, R., Lee, J., Bluestone, J. A., Anderson, M., & Herold, K. C. (2018). Collateral damage: Insulin-dependent diabetes induced with checkpoint inhibitors. *Diabetes, 67*(8), 1471–1480.

[5] Tsang, V. H. M., McGrath, R. T., Clifton-Bligh, R. J., Scolyer, R. A., Jakrot, V., Guminski, A. D., Long, G. V., & Menzies, A. M. (2019). Checkpoint inhibitor-associated autoimmune diabetes is distinct from type 1 diabetes. *Journal of Clinical Endocrinology & Metabolism, 104*(11), 5499–5506.

[6] Wu, L., Tsang, V. H. M., Sasson, S. C., Menzies, A. M., Carlino, M. S., Brown, D. A., Clifton-Bligh, R., & Gunton, J. E. (2021). Unravelling checkpoint inhibitor associated autoimmune diabetes: From bench to bedside. *Frontiers in Endocrinology, 12*, 764138.

[7] Wu, L., Tsang, V., Menzies, A. M., Sasson, S. C., Carlino, M. S., Brown, D. A., Clifton-Bligh, R., & Gunton, J. E. (2023). Risk factors and characteristics of checkpoint inhibitor-associated autoimmune diabetes mellitus (CIADM): A systematic review and delineation from type 1 diabetes. *Diabetes Care, 46*(6), 1292–1299.

[8] Byun, D. J., Braunstein, R., Flynn, J., Zheng, J., Lefkowitz, R. A., Kanbour, S., & Girotra, M. (2020). Immune checkpoint inhibitor-associated diabetes: A single-institution experience. *Diabetes Care, 43*(12), 3106–3109.

[9] Wu, L., Carlino, M. S., Brown, D. A., Long, G. V., Clifton-Bligh, R., Mellor, R., Moore, K., Sasson, S. C., Menzies, A. M., Tsang, V., & Gunton, J. E. (2024). Checkpoint inhibitor-associated autoimmune

diabetes mellitus is characterized by C-peptide loss and pancreatic atrophy. *Journal of Clinical Endocrinology & Metabolism, 109*(5), 1301–1307.

[10] Wei, H.-H., Lai, Y.-C., Lin, G., Lin, C.-W., Chang, Y.-C., Chang, J. W.-C., Liou, M.-J., & Chen, I.-W. (2024). Distinct changes to pancreatic volume rather than pancreatic autoantibody positivity: Insights into immune checkpoint inhibitors induced diabetes mellitus. *Diabetology & Metabolic Syndrome, 16*(1), 26.

[11] Wu, L., Wentworth, J. M., Liddle, C., Fewings, N., Carlino, M., Brown, D. A., Clifton-Bligh, R., Long, G. V., Scolyer, R. A., Norris, N., Sasson, S. C., Tsang, V. H. M., Menzies, A. M., & Gunton, J. E. (2026). Pancreatic volume and immune biomarkers predict checkpoint inhibitor–associated autoimmune diabetes in humans. *Journal of Clinical Investigation, 136*(4), e192938.

[12] Virostko, J., Williams, J. M., Hilmes, M. A., Bowman, C., Wright, J. J., Du, L., Kang, H., Russell, W. E., Powers, A. C., & Moore, D. J. (2019). Pancreas volume declines during the first year after diagnosis of type 1 diabetes and exhibits altered diffusion at disease onset. *Diabetes Care, 42*(2), 248–257.

[13] Virostko, J., Wright, J. J., Williams, J. M., Hilmes, M. A., Triolo, T. M., Broncucia, H., Du, L., Kang, H., Nallaparaju, S., Valencia, L. G., Reyes, D., Hammel, B., Russell, W. E., Philipson, L. H., Waibel, M., Kay, T. W. H., Thomas, H. E., Greeley, S. A. W., Steck, A. K., . . . Moore, D. J. (2024). Longitudinal assessment of pancreas volume by MRI predicts progression to stage 3 type 1 diabetes. *Diabetes Care, 47*(3), 393–400.

[14] Wasserthal, J., Breit, H.-C., Meyer, M. T., Pradella, M., Hinck, D., Sauter, A. W., Heye, T., Boll, D. T., Cyriac, J., Yang, S., Bach, M., & Segeroth, M. (2023). TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. *Radiology: Artificial Intelligence, 5*(5), e230024.

[15] Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods, 18*(2), 203–211.

[16] Lim, S.-H., Kim, Y. J., Park, Y.-H., Kim, D., Kim, K. G., & Lee, D.-H. (2022). Automated pancreas segmentation and volumetry using deep neural network on computed tomography. *Scientific Reports, 12*(1), 4075.

[17] von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet, 370*(9596), 1453–1457.

[18] Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A.-L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., . . . Logullo, P. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378.

[19] Saisho, Y., Butler, A. E., Meier, J. J., Monchamp, T., Allen-Auerbach, M., Rizza, R. A., & Butler, P. C. (2007). Pancreas volumes in humans from birth to age one hundred taking into account sex, obesity, and presence of type-2 diabetes. *Clinical Anatomy, 20*(8), 933–942.

[20] Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163.

[21] Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307–310.

[22] Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38*(4), 963–974.

[23] Heagerty, P. J., Lumley, T., & Pepe, M. S. (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics, 56*(2), 337–344.

[24] Vickers, A. J., & Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. *Medical Decision Making, 26*(6), 565–574.

[25] Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Jr., Martin, G. P., Reitsma, J. B., Moons, K. G. M., Collins, G., & van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ, 368*, m441.