Supplementary Section:

Data Source:

The data utilized in this study were obtained from the TriNetX (Cambridge, MA) Research Network. This federated health research platform offers real-time access to electronic health record (EHR) data collected from 56 healthcare organizations (HCOs), predominantly located in the United States. All data is de-identified and aggregated directly from the EHR of participating HCOs. This data includes information such as diagnoses, laboratory results, medications, procedures, and genomic information. The HCOs care for a diverse patient population with varying ethnicities, age groups, geographical locations, and income levels. Both the patients and the HCOs contributing the data remain anonymous; thus, no Institutional Review Board approval was required for the de-identified data.

Standardizing the Terminology and Data Quality Check:

TriNetX utilizes robust processes to map data from various sources to a standardized model effectively. This includes mapping codes and extracting information from clinical documents using natural language processing. TriNetX can then stratify the data based on a set of clinical terminologies, and enforce required fields (e.g., patient identifiers) while rejecting patient records that are missing the required data. Referential integrity checks are performed, and the ongoing monitoring of data volumes confirms the validity of the data over time.

Clinical Fact and Coding System to Present Data:

TriNetX utilizes various techniques to ensure data quality and consistency. The data are extensively mapped to a standard model within TriNetX by a team of informaticists, who convert the local codes used by data providers into master terminology. Laboratory test results, vitals, and other findings are coded using Logical Observation Identifier Names and Codes (LOINC), while demographic information follows the Health Level 7 (HL7) version 3 (administrative standards). Diagnoses are represented by the International Classification of Diseases, Ninth and Tenth Revisions, Clinical Modification (ICD-9-CM and ICD-10-CM). Procedure data are coded using either the

Current Procedural Terminology (CPT) or the ICD-9 and ICD-10 Procedure Coding System, depending on the coding system utilized by the HCO.

Definition of Study Cohort: We defined Coeliac disease using the following inclusion criteria: ICD-9 code 579.0 (Celiac Disease) and ICD-10 code K90.0 (Celiac disease). We defined chronic pancreatitis, acute pancreatitis, and the subtypes utilizing the following inclusion criteria: ICD-9 code of 577.1 (Chronic Pancreatitis), ICD-10 code of K86.1 (Other Chronic Pancreatitis); ICD-9 code of 577.0 (Acute Pancreatitis), ICD-10 code of K85.9 (Acute Pancreatitis); ICD-10 code of K85.2 (Alcohol Induced Acute Pancreatitis); ICD-10 code of K85.1 (Biliary Acute Pancreatitis); ICD-10 code of K85.0 (Idiopathic acute pancreatitis).

Selection of Patients with Non-CD Cohort: To establish a control group, we excluded patients with a diagnosis of celiac disease, as indicated by the aforementioned codes.

Supplementary Table 1: Laboratory characteristics of Patients with CD and without CD Before and After Propensity Score Matching

Variables	**Before Propensity Matching**			**After Propensity Matching**		
	CD (n= 160228)	**Non-CD** (n=250725)	**SMD**	**CD** (n=137685)	**Non-CD** (n=137685)	**SMD**
Hematological Profile, Mean (SD)						
Hemoglobin (g/dL)	13.1(1.8)	12.8(2.1)	0.1590	13.1(1.9)	12.9(2.0)	0.0975
Hematocrit (%)	39.3(5.2)	38.6(6.0)	0.1245	39.3(5.3)	38.8(5.9)	0.0995
Creatinine (mg/dL)	0.816(1.2)	1(2.0)	0.1138	0.834(1.3)	0.946(2.0)	0.0670
Platelets [K/uL] in blood	269(88)	254(89.9)	0.1581	266(87.6)	254(89.9)	0.1333
Metabolic Profile, Mean (SD)						
Glucose (mg/dL)	104(49.5)	105(39.2)	0.0184	104(48)	104(37.7)	0.0046
Hemoglobin A1c (mmol/L)	6.47(2.1)	6.23(1.8)	0.1224	6.35(2.0)	6.22(1.9)	0.0696

Bicarbonate (mEq/L)	25.4(3.3)	25.8(3.2)	0.1174	25.5(3.2)	25.7(3.2)	0.0556
Calcium (mg/dL)	9.27(0.6)	9.27(0.6)	0.0033	9.26(0.6)	9.29(0.6)	0.0445
Albumin (g/dL)	4.14(0.5)	4.02(0.6)	0.2217	4.13(0.5)	4.07(0.6)	0.1043
Protein (g/dL)	6.97(1.1)	7.07(1.2)	0.0938	6.97(1.1)	7.02(1.2)	0.0490
Folate (ng/mL)	13.4(10.0)	13.5(8.4)	0.0127	56.2(136)	13.8(8.8)	0.0432
Lactate (mg/dL)	1.59(3.5)	1.47(2.2)	0.0400	1.58(3.6)	1.54(2.5)	0.0097
Lipid Profile, Mean (SD)						
Cholesterol (mg/dL)	178(43.7)	179(44.3)	0.0363	178(43.9)	180(43.3)	0.0303
HDL (mg/dL)	51.7(19.6)	51.3(19.5)	0.0222	51.7(19.8)	51.8(19.6)	0.0035
LDL (mg/dL)	102(35.1)	102(36.5)	0.0066	103(35.3)	103(35.7)	0.0005
Triglycerides (mg/dL)	120(96.8)	130(95.9)	0.1004	121(95.9)	129(98.8)	0.0844
Liver Profile, Mean (SD)						
INR	1.33(2.8)	1.26(1.2)	0.0326	1.32(2.7)	1.26(1.1)	0.0256
Prothrobin Time (s)	12.8(6.6)	13.6(6.2)	0.1194	12.8(6.5)	13.7(5.9)	0.1448
APTT(s)	30(11.2)	32.7(14.1)	0.2121	30(11.2)	32.9(14.3)	0.2231
Bilirubin total (mg/dL)	0.557(0.8)	0.593(0.9)	0.0416	0.56(0.8)	0.61(1)	0.0556
ALT (U/L)	29.2 (65.4)	31.8 (72.9)	0.0371	29.3 (67.3)	33.8 (89)	0.0571
AST (U/L)	28.8 (69.3)	30.7 (75.7)	0.0265	28.8(69.3)	31.7(92.8)	0.0366
GGT(U/L)	54.3(132)	93.4(207)	0.2246	65.7(72.7)	83.1(180)	0.1686
ALP (IU/L)	99.2(91.2)	88.4(66.9)	0.1341	93.7(73.8)	90.7(75.4)	0.0401
LDH (IU/L)	243(335)	285(455)	0.1074	241(326)	282(322)	0.1257
Inflammatory Markers, Mean (SD)						
EST (mm/h)	16.3(18.4)	23.8(1.2)	0.3525	16.6(18.6)	20.5(21.8)	0.1908
C- reactive protein (mg/L)	11.4(29.9)	17.1(37.9)	0.1650	11.4(29.9)	16.8(37.8)	0.1569

Digestive Enzymes, Mean (SD)						
Lipase (U/L)	74.5(762)	71.3(304)	0.0055	72.8(747)	70.3(369)	0.0043
Amylase (U/L)	66.1(93.4)	68.1(88.1)	0.0217	65.7(72.7)	65.9(94.5)	0.0019
Gliadin Antibodies [Units/volume], Mean (SD)						
Gliadin peptide IgA Ab	17.7(22.7)	6.53(10.5)	0.6309	17.4(22.5)	6.55(11.4)	0.6098
Gliadin peptide IgG Ab	25.7(28.7)	7.6(13.3)	0.8093	25.3(28.5)	9.5(17.1)	0.6717
Gliadin IgG Ab	14.8(17.9)	4.77(6.1)	0.7485	14.7(18)	5.42(7.5)	0.6714
Other, Mean (SD)						
Urate (mg/dL)	5.18(2.0)	5.56(2.3)	0.1760	5.22(2.0)	5.42(2.2)	0.0963

Abbreviations: CD, celiac disease; INR, International Normalized Ratio; Ab, antibodies; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; APTT, activated partial thromboplastin time; EST, erythrocyte sedimentation rate; GGT, gamma glutamyl transferase; LDL, low density lipoprotein;LDH, lactate dehydrogenase HDL, high-density lipoprotein; SD, standard deviation; SMD, standard mean difference.