Supplementary Tables

Supplementary Table 1 List of Codes (Finnish Version) for Gastrointestinal Bleedings

ICD-10 Codes	Description
I850	Esophageal varices
K226	Gastro-esophageal laceration syndrome (Mallory–Weiss syndrome)
K250	Acute gastric ulcer with hemorrhage (peptic ulcer disease (pud), peptic ulcer or stomach ulcer)
K252	Acute gastric ulcer with both hemorrhage and perforation (pud)
K254	Chronic or unspecified gastric ulcer with hemorrhage (pud)
K256	Chronic or unspecified gastric ulcer with both hemorrhage and perforation (pud)
K260	Acute duodenal ulcer with hemorrhage (pud)
K262	Acute duodenal ulcer with both hemorrhage and perforation (pud)
K264	Chronic or unspecified duodenal ulcer with hemorrhage (pud)
K266	Chronic or unspecified duodenal ulcer with both hemorrhage and perforation (pud)
K270	Acute peptic ulcer, site unspecified, with hemorrhage (pud)
K272	Acute peptic ulcer, site unspecified, with both hemorrhage and perforation (pud)
K274	Chronic or unspecified peptic ulcer, site unspecified, with hemorrhage (pud)
K276	Chronic or unspecified peptic ulcer, site unspecified, with both hemorrhage and perforation (pud)
K290	Acute bleeding gastric ulcer
K661	Hemoperitoneum
K920	Hematemesis
K22.8	Other specified diseases of esophagus (Includes: Bleeding in the esophagus UNS), more precisely defined esophageal bleeding
K280	Acute gastrojejunal ulcer with hemorrhage
K282	Acute gastrojejunal ulcer with both hemorrhage and perforation
K284	Chronic or unspecified gastrojejunal ulcer with hemorrhage
K286	Chronic or unspecified gastrojejunal ulcer with both hemorrhage and perforation
K5280 #	Antibiotic-associated [hemorrhagic] colitis
K625	Hemorrhage of anus and rectum
K921	Melena
K922	Gastrointestinal hemorrhage, unspecified (Stomach bleeding UNS / Intestinal bleeding UNS)

ICD-9 Codes	Description
4560A	Esophageal varices with bleeding
5307A	Mallory–Weiss syndrome
531*B	Ulcer with bleeding
532*B	Ulcer with bleeding
5780A	Hematemesis
5781A	Melena
5621C	Diverticulitis coli with hemorrhage
5781B	Hemorrhage ex ano

Abbreviations: Pud: Peptic Ulcer Disease; UNS: Unspecified

Supplementary Table 2 List of Codes (Finnish version) for comorbidities using Hospital discharge register, Cause of death register, Drug Reimbursement register, and Prescription register.

Diagnosis	ICD Codes
Any cancer	ICD 9 - 14-16, 170-172, 174-179, 175-179, 18-19, 200-208
	ICD 10 - C0-C3, C40-C43, C45-C49, C5-C9
Any mental disease	ICD 9 – 29, 30-31
	ICD 10 - F
	Drug Reimbursement Register – 112
	Prescription Register - N05A
Stroke, including sub-arachnoid hemorrhage	ICD 9 - 430, 431, 4330A, 4331A, 4339A, 4340A, 4341A, 4349A, 436
	ICD 10 - I60, I61, I63, I64; excluding I636
Venous thromboembolism	ICD 9 – 415, 451, 6713, 6714, 6732; excluding 4510
	ICD 10 - I26, I80, O871, O882; excluding I800
Ischaemic heart disease	ICD 9 - 410-414, 798; excluding 7980A
	ICD 10 - I20-I25, I21-I25, I46, R96, R98
	Drug Reimbursement Register – 206
Valvular heart disease	ICD 9 - 394-396, 4240-4243, 7460, 7461, 7464-7466
	ICD 10 - I05-I08, I34-I37
Atrial fibrillation	ICD 9 - 4273
	ICD 10 - I48
	Drug Reimbursement Register – 207, 351, I48
Peripheral artery disease	ICD 9 – 4402, 4439
	ICD 10 - E105, E115, E125, E135, E145, I702, I739
Heart failure	ICD 9 - 4029B, 404, 4148, 428
	ICD 10 - I50, I110, I130, I132
	Drug Reimbursement Register – 201
	Prescription Register - C03CA01, C03EB01
High blood pressure	ICD 9 – 401
	ICD 10 - I10
	Drug Reimbursement Register – 205
Osteoarthritis	ICD 9 – 715
	ICD 10 - M15-M19
Inflammatory connective tissue diseases including rheumatoid arthritis	ICD 9 – 710, 714, 420, 725
	ICD 10 - M05-M09, M10-M13, M32, M33, M45
	Drug Reimbursement Register – 202
Diabetes; excluding gestational diabetes	ICD 9 - 250
	ICD 10 - E10-E14
	Drug Reimbursement Register – 103
	Prescription Register - A10
Chronic obstructive pulmonary disease	ICD 9 – 491, 492
	ICD 10 - J43, J44
Asthma	ICD 9 – 493; excluding 4939C
	ICD 10 - J45, J46
	Drug Reimbursement Register – 203
	Prescription Register - R03BA, R03BC, R03DC, R03AK
Liver disorders	ICD 9 – 570-573

Diagnosis	ICD Codes
Precursors of gastrointestinal Bleeding (w/o bleeding: Gastritis, Colitis, intestinal diseases, Crohn's disease, esophageal varices, esophagitis, hemorrhoids)	ICD 10 - K70-K77 ICD 9 - 5302A-5302B; 535; 555; 556; 557; 4561A; 530; 455; 6718 ICD 10 - K22.1 #; K29.1; K29.2; K29.3; K29.4; K29.5, K29.6; K29.7; K29.8; K29.9; K50; K51; K55; I85.9; K20; K20 #; I84; O22.4; O87.2

Abbreviations: ATC: Anatomical Therapeutic Chemical Classification System; DDD: Defined Daily Dose; ICD-9: International Classification of Diseases, Ninth Revision; ICD-10: International Statistical Classification of Diseases and Related Health Problems, Tenth Revision; WHO: World Health Organization

Supplementary Table 3 Estimated coefficients for least absolute shrinkage and selection operator using cox regression for the risk factors of major gastrointestinal bleeding. Negative coefficients show decreased risk, positive coefficients show increased risk, and zero coefficients are as the least important variables in the model to predict gastrointestinal bleeding

Variable	Coefficient
Baseline age	0.04
Gender	
Male	Ref.
Female	-0.16
Year of enrollment	
1987	Ref.
1992	0
1997	0.20
2002	0
2007	-0.10
Region[1]	
West	Ref.
East	-0.02
Marital status	
Married, cohabiting, or registered partnership	Ref.
Single, separated, divorced, or widow	0.05
Occupation	
Non-manual (office/studying)	Ref.
Manual (factory/construction/farming/forestry)	0.00
Family/housewife/pensioner	0.14
Unemployed	0.36
Education in Tertiles[2]	
Low	Ref.
Moderate	-0.16
High	0
Smoking	
Never	Ref.
Ex-Smoker	0.07
Smoker	0.12
Alcohol consumption in grams per week	

Variable	Coefficient
Non-drinker	Ref.
Mild (0.1g - 36.9g)	-0.04
Moderate (37g – 86.9g)	-0.17
Heavy (≥87g)	0.02
Coffee consumption per day	
0 cups	Ref.
1 – 5 cups	0
6 – 10 cups	-0.01
>10 cups	-0.26
Physical activity	
Minimal - mild	Ref.
Moderate to heavy	-0.09
Body mass index stnzd.	0.07
Waist hip ratio stnzd.	0.19
Systolic blood pressure stnzd.	-0.11
Diastolic blood pressure stnzd.	0
Total cholesterol	-0.03
High density lipoprotein stnzd.	0
Gamma-glutamyl transferase stnzd.	0.05
Precursors of GIB[3]	
0	Ref.
>=1	0.61
Any cancer	0.36
Any psychiatric disorders	0.26
Stroke including SAH	0.22
Venous thromboembolism	0
Ischaemic heart disease	0.14
Valvular heart disease	0.25
Atrial fibrillation	0.03
Peripheral artery disease	0.28
Heart failure	0.40
High blood pressure	-0.01
Osteoarthritis	-0.73
Inflammatory connective tissue diseases	0
Diabetes[4]	-0.15
Chronic obstructive pulmonary disease	0.17

Variable	Coefficient
Asthma	-0.09
Liver disorders	1.18
Anemia	0.25

Ref.: - Reference; GIB: Gastrointestinal bleeding; SAH: Sub-arachnoid hemorrhage
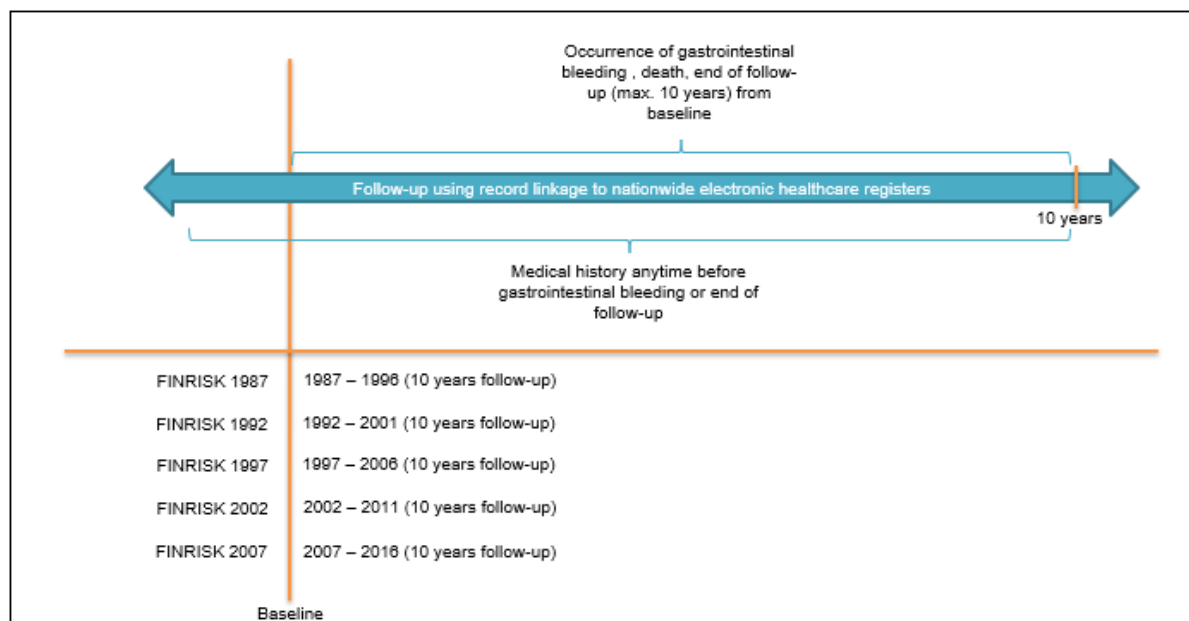
[1] West Finland includes Turku and Loimaa as well as Helsinki and Vantaa. East Finland includes North Karelia, North Savo, Oulu, and Lapland.

[2] Educational tertiles were calculated according to years of education and were specific to the birth cohorts.
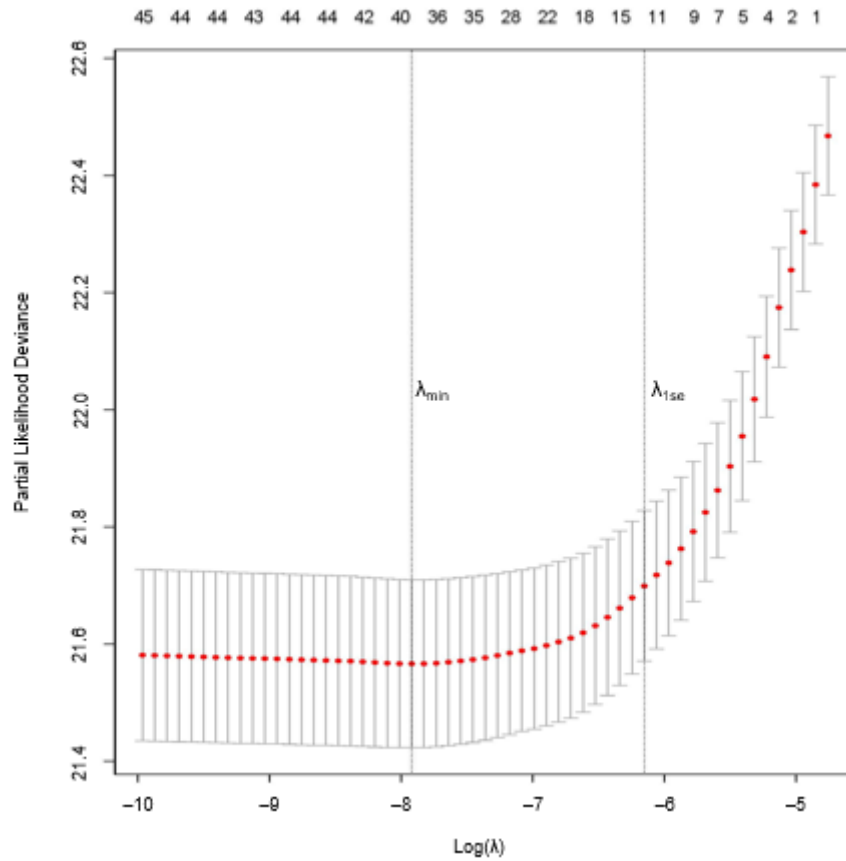
[3] Includes any diagnosis of gastritis, intestinal diseases, esophageal varices, esophagitis, hemorrhoids, colitis, or Crohn's disease – without gastrointestinal bleeding.

[4] Excluding gestational diabetes

Supplementary Figures

Supplementary Figure 1 Study design overview. Baseline data collected at enrollment includes demographics, socioeconomic factors, lifestyle factors, blood pressure measurements, and laboratory measurements. Occurrence of gastrointestinal bleedings (GIB) was observed within 10 years after enrollment. Follow-up was censored at GIB event, death or end of follow-up and medical history was observed any time before censoring.

Supplementary Figure 2 Cross validation plot for the penalty term generated from the least absolute shrinkage and selection operator Cox regression. The top x-axis shows the number of variables in the model from 1-46. The bottom x-axis shows log (λ) values for the dashed lines corresponding to the minimum λ value (λ_{min} - left dashed line) and λ within one standard deviation (λ_{1se} - right dashed line). λ_{min} is the value for which the model with the respective number of variables has the lowest partial likelihood deviance i.e., smaller the deviance – better the fit. This deviance has a certain variability as shown using grey error bars to every red point. λ_{min} and λ_{1se} gives us a range of variables (12-39) which balances between model accuracy and model parsimony.