

**PREMATURE MORTALITY FROM
CARDIOVASCULAR DISEASE: GLOBAL
BURDEN AND COUNTRY-SPECIFIC ESTIMATE
THROUGH REVIEW STUDY, TREND ANALYSIS,
PREDICTION MODELLING, AND CAUSAL
INFERENCE STUDY**

WAN SHAKIRA BINTI RODZLAN HASANI

UNIVERSITI SAINS MALAYSIA

2024

**PREMATURE MORTALITY FROM
CARDIOVASCULAR DISEASE: GLOBAL
BURDEN AND COUNTRY-SPECIFIC ESTIMATE
THROUGH REVIEW STUDY, TREND ANALYSIS,
PREDICTION MODELLING, AND CAUSAL
INFERENCE STUDY**

by

WAN SHAKIRA BINTI RODZLAN HASANI

**Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy**

DECEMBER 2024

ACKNOWLEDGEMENT

I extend my deepest gratitude to my main supervisor, Professor Dr. Kamarul Imran Musa, for his invaluable academic support, guidance, and encouragement throughout the entirety of this PhD journey. His flexibility, constructive feedback, and understanding have made my PhD experience smooth. I am truly grateful for his brilliant ideas in designing my thesis. My heartfelt appreciation also goes to my co-supervisor, Associate Professor Dr. Yee Cheng Kueh, for their consistent and dedicated efforts in reviewing every page of this thesis and draft manuscripts. I would like to express my gratitude to the Ministry of Health Malaysia for providing full-time study leave and scholarship, which made this journey possible. Special thanks to my research team members from the National Institutes of Health, Ministry of Health Malaysia, especially Datuk Dr. Nor Asiah Muhamad and Dr. Azahadi Omar, as well as the team from the Institute for Public Health (Dr Noor Ani Ahmad, Dr Muhammad Fadhli and Dr Shubash Shander). I am deeply indebted to Dr. Tengku Muhammad Hanis, my senior PhD coursemate, whose guidance, particularly in the analysis part, from the beginning of the PhD journey until the end. I have learned a lot from him.

To my parents (Rodzlan Hasani Muhammed and Wan Sharipah Nik Sof), thank you for your immense support, understanding, and never-ending dua throughout my journey. To my beloved husband, Mohd Azamli, thank you for your unwavering support. Your presence has made this three-year journey delightful, wonderful, and filled with happiness. To my son, Thalhah, you have been my source of strength and positivity. Despite the recent diagnosis of a genetic disorder, you remain my special child, and I will forever cherish you. You are undeniably a precious gift from Allah.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	ii
TABLE OF CONTENTS.....	iii
LIST OF TABLES	xi
LIST OF FIGURES	xv
LIST OF SYMBOLS	xix
LIST OF ABBREVIATIONS	xx
LIST OF APPENDICES	xxiii
ABSTRAK	xxiv
ABSTRACT	xxvii
CHAPTER 1 INTRODUCTION.....	1
1.1 Background of the study	1
1.2 Problem statement	3
1.3 Research questions	7
1.4 Research objectives	8
1.4.1 General objective.....	8
1.4.2 Specific objectives.....	8
1.5 Overview of studies.....	9
1.5.1 Study 1: Bibliometric analysis on premature CVD mortality	10
1.5.2 Study 2a: Systematic review on years of life lost from CVD mortality	10
1.5.3 Study 2b: Meta analysis of ASMR from premature CVD Mortality.....	11
1.5.4 Study 3: Trend of ASMR from CVD in Malaysia	11
1.5.5 Study 4: Prognostic Factors for Premature CVD Mortality in Malaysia	12
1.5.6 Study 5: Constructing DAG model for Premature CVD Mortality.....	13

CHAPTER 2	LITERATURE REVIEW	14
2.1	Introduction	14
2.2	Cardiovascular diseases.....	14
2.2.1	Definition CVD	14
2.2.2	Main type of CVD.....	15
2.2.2(a)	Ischemic heart disease	15
2.2.2(b)	Stroke	16
2.2.3	Classification of CVD (codes)	17
2.2.4	Epidemiology of CVD	19
2.2.4(a)	Incidence, burden, and trend of CVD	19
2.2.4(b)	The global burden of IHD.....	19
2.2.4(c)	The global burden of stroke	20
2.2.4(d)	Epidemiological transitions	20
2.3	Premature mortality	20
2.3.1	Definition	20
2.3.2	Years of life lost	21
2.3.3	Age standardized mortality rate	22
2.3.4	Global burden of Premature CVD mortality	24
2.3.5	Country specific burden of premature CVD mortality.....	25
2.4	CVD mortality risk factors	26
2.4.1	Non-modifiable risk factors	26
2.4.2	Modifiable risk factors	29
2.4.3	Social determinants of health	35
2.5	CVD risk prediction models.....	37
2.6	Review methodologies	38
2.6.1	Bibliometric analysis.....	39
2.6.2	Systematic review and meta-analysis.....	41

2.6.3	Comparison three type of review studies	44
2.7	Trend analysis	45
2.7.1	Definition and type of trend analysis	45
2.7.2	Joinpoint regression.....	46
2.8	Survival modelling study	48
2.8.1	Types of survival analysis	49
2.8.1(a)	Non-parametric survival analysis	50
2.8.1(b)	Semi-parametric survival analysis.....	51
2.8.1(c)	Parametric survival analysis	52
2.8.2	Imbalanced data, survival modeling, and resampling techniques.....	55
2.9	Causal inference	56
2.9.1	Causal inference overview	56
2.9.2	Directed acyclic graph (DAG)	58
2.9.3	Causal inference framework.....	63
2.10	Delphi Method.....	65
2.10.1	Delphi Method overview.....	65
2.10.2	Fuzzy Delphi Method.....	65
CHAPTER 3 METHODOLOGY.....		67
3.1	Introduction	67
3.2	Study 1: Bibliometric Analysis of Premature Mortality	69
3.2.1	Data source and literature inclusion criteria.....	69
3.2.2	Data extraction and collection.....	70
3.2.3	Data analysis and software	71
3.3	Study 2a: Systematic review of YLL of premature CVD mortality	73
3.3.1	Study protocol	73
3.3.2	Search strategy	73
3.3.3	Study selection	76

3.3.4	Data extraction and management	79
3.3.5	Quality assessment	80
3.3.6	Research team (screening and data extraction)	81
3.4	Study 2b: Meta-analysis of ASMR for Premature CVD Mortality	85
3.4.1	Protocol and registration	85
3.4.2	Search strategy	85
3.4.3	Inclusion and exclusion criteria.....	90
3.4.4	Screening process	90
3.4.5	Data extraction and management	91
3.4.6	Quality assessment	92
3.4.7	Statistical Analysis	95
3.4.8	Assessment of heterogeneity	97
3.4.9	Assessment of reporting biases	97
3.4.10	Sensitivity analysis	98
3.4.11	Research team (screening and data extraction)	99
3.5	Study 3: Trend of ASMR from CVD in Malaysia	100
3.5.1	Study design and target population	100
3.5.2	Source of data.....	100
3.5.3	Statistical analysis	101
3.5.3(a)	Age standardized mortality rate calculation	101
3.5.3(b)	Joinpoint regression analysis	102
3.5.4	Data source approval.....	105
3.6	Study 4: Prognostic factors for premature CVD mortality in Malaysia.....	105
3.6.1	Study design and subjects	106
3.6.2	Data sources	107
3.6.3	Baseline data collection.....	107
3.6.4	Definition of terms and study variables	108

3.6.4(a)	Outcome variable.....	108
3.6.4(b)	Prognostic variables.....	108
3.6.5	Sample size calculation.....	113
3.6.6	Statistical analyses.....	116
3.6.6(a)	Incidence rate.....	116
3.6.6(b)	Semi-parametric survival analysis	120
3.6.6(c)	Parametric survival analysis	121
3.6.6(d)	Sensitivity analysis	124
3.7	Study 5: Constructing Causal Pathways for Premature CVD Mortality	125
3.7.1	Overview study stage	125
3.7.2	Stage 1: Evidence synthesis for constructing DAGs (ESC-DAG).....	128
3.7.3	Stage 2: Delphi method for expert consensus	130
3.7.3(a)	First round interview.....	130
3.7.3(b)	Second round interview	130
3.7.3(c)	Fuzzy Delphi method.....	131
3.7.4	Stage 3: Testing DAG model with available data	134
3.8	Ethical approval.....	136
CHAPTER 4 RESULTS AND DISCUSSION.....		137
4.1	Introduction	137
4.2	Study 1: Bibliometric Analysis of Premature Mortality	137
4.2.1	Results	137
4.2.1(a)	Publication by year	138
4.2.1(b)	Publication by country	139
4.2.1(c)	Publication by journal.....	141
4.2.1(d)	Publication by authors	142
4.2.1(e)	Keywords analysis	144
4.2.2	Discussion	151

4.2.3	Strengths and limitations.....	155
4.3	Study 2a: Systematic review of YLL of premature CVD mortality	156
4.3.1	Results	156
4.3.1(a)	Quality assessment.....	156
4.3.1(b)	Characteristic of included studies.....	156
4.3.1(c)	Years of potential life lost (YPLL) from CVD.....	165
4.3.1(d)	Standard expected years of life lost (SEYLL) from CVD.....	168
4.3.1(e)	The pattern of years of life lost from CVD.....	171
4.3.2	Discussion	176
4.3.3	Strengths and limitations.....	179
4.4	Study 2b: Meta-analysis of ASMR for premature CVD mortality	181
4.4.1	Results	181
4.4.1(a)	Study characteristics	181
4.4.1(b)	Quality assessment.....	186
4.4.1(c)	Overall ASMR estimates	186
4.4.1(d)	Subgroup analysis by CVD types	187
4.4.1(e)	Subgroup analysis by sex.....	189
4.4.1(f)	Subgroup analysis by income country level	194
4.4.1(g)	Subgroup analysis by study time	197
4.4.1(h)	Sensitivity analysis – removing outliers.....	201
4.4.1(i)	Sensitivity analysis – stratified by age thresholds	204
4.4.1(j)	Meta-regression model	206
4.4.1(k)	Publication’s bias.....	207
4.4.2	Discussion	208
4.4.3	Strengths and limitations.....	213
4.5	Study 3: Trend of ASMR from CVD in Malaysia	214
4.5.1	Results	214

4.5.1(a)	Age standardized mortality rate.....	215
4.5.1(b)	Joinpoint regression model	218
4.5.2	Discussion	226
4.5.3	Strengths and limitations.....	230
4.6	Study 4: Factors Associated with Premature CVD Mortality in Malaysia ..	231
4.6.1	Results	231
4.6.1(a)	Semi-parametric survival analysis model.....	237
4.6.1(b)	Parametric survival analysis model	247
4.6.1(c)	Sensitivity analysis	251
4.6.2	Discussion	257
4.6.3	Limitations and strengths	265
4.7	Study 5: Constructing Causal Pathways for Premature CVD Mortality	267
4.7.1	Results stage 1 (DAG model from ESC-DAG protocol)	267
4.7.1(a)	Mapping	267
4.7.1(b)	Translation	271
4.7.1(c)	Integration.....	274
4.7.2	Results stage 2 (Expert Consensus through Delphi Method).....	276
4.7.2(a)	First round interview.....	276
4.7.2(b)	Second round interview	278
4.7.2(c)	Final consensus from Fuzzy Delphi Method	278
4.7.2(d)	Suggested adjustment set.....	286
4.7.3	Results stage 3 (Application of DAG model with available data).....	289
4.7.3(a)	Model from Expert Ranked	293
4.7.3(b)	Model accuracy from available data.....	295
4.7.4	Discussion	298
4.7.5	Strengths and limitations.....	301

CHAPTER 5	CONCLUSION AND RECOMMENDATIONS	303
5.1	Overview	303
5.2	Conclusion and recommendations for study 1	303
5.3	Conclusion and recommendations for study 2a	304
5.4	Conclusion and recommendations for study 2b	304
5.5	Conclusion and recommendations for study 3	305
5.6	Conclusion and recommendations for study 4	306
5.7	Conclusion and recommendations for study 5	306
5.8	Overall summary	307
5.9	Study contributions	308
REFERENCES.....		311
APPENDICES		
LIST OF PUBLICATIONS		

LIST OF TABLES

	Page
Table 2.1 The International Classification of Diseases, 10th revision (ICD-10) for main CVD	18
Table 2.2 Studies reporting non-modifiable risk factors for CVD mortality	27
Table 2.3 Studies reporting modifiable risk factors for CVD mortality	30
Table 3.1 Search terms for bibliometric study on premature mortality	70
Table 3.2 Search terms for systematic review year of life lost from premature CVD mortality.....	74
Table 3.3 Formulas for Years of Potential Life Lost (YPLL) and Standard Expected Years of Life Lost (SEYLL)	78
Table 3.4 Data extraction form for systematic review of years of life lost from premature CVD mortality.....	82
Table 3.5 Newcastle-Ottawa Scale for quality assessment	83
Table 3.6 Search terms for systematic review of ASMR from premature CVD Mortality	87
Table 3.7 The adapted version of Newcastle-Ottawa Scale for this study (Systematic review and Meta-analysis of ASMR for Premature CVD Mortality).....	93
Table 3.8 List of prognostic variables for premature CVD mortality study in Malaysia	110
Table 3.9 Sample size calculation for survival analysis for premature CVD mortality risk profile	114
Table 3.10 Calculation of age standardized incidence rate: Premature CVD mortality (ICD10: I00-I99) aged 30-69 years	117
Table 3.11 Calculation of age standardized incidence rate: Premature CVD mortality (ICD10: I00-I99) aged 30-69 years among male	118

Table 3.12	Calculation of age standardized incidence rate: Premature CVD mortality (ICD10: I00-I99) aged 30-69 years among female	119
Table 3.13	Triangular Fuzzy number from 7 Likert-scale	132
Table 4.1	The top 10 countries contributed to the premature mortality studies according to total publication and total citation	140
Table 4.2	The core 10 journals published the premature mortality articles with the number of publications	142
Table 4.3	The top 10 most productive authors of premature mortality study ..	143
Table 4.4	The frequent keywords in the premature mortality studies	145
Table 4.5	Characteristics of selected studies; systematic review of years of life lost from cardiovascular diseases	160
Table 4.6	Summary of total years of potential life lost (YPLL) from CVD mortality according to study characteristics	166
Table 4.7	Summary of total standard expected years of life lost (SEYLL), from CVD mortality method from GBD study (1996) according to study characteristics	169
Table 4.8	Characteristics of included studies; systematic review and meta-analysis of ASMR from premature CVD mortality	183
Table 4.9	Quality assessment of 15 selected studies using an adapted version of the Newcastle-Ottawa Scale for cross-sectional studies	185
Table 4.10	Summary of pooled ASMR per 100,000 population for premature cardiovascular mortality; Meta-analysis using random effect model	187
Table 4.11	Sensitivity analysis of pooled ASMR per 100,000 population from premature CVD mortality by excluding the studies outlier	202
Table 4.12	Sensitivity analysis of pooled ASMR per 100,000 population from premature CVD mortality by difference age threshold	205
Table 4.13	Multivariable meta-regression model of log ASMR estimates from premature CVD mortality	207

Table 4.14	Population structure and age standardized mortality rate from CVD in Malaysia from 2010-2021	216
Table 4.15	Joinpoint regression model of age standardized mortality rate from CVD in Malaysia, 2010-2021	220
Table 4.16	Characteristics of Malaysia adults aged 18-69 years at baseline sampled from National Health and Morbidity Surveys (2006, 2011, 2015) and 15 years follow-up of premature cardiovascular mortality	233
Table 4.17	Baseline characteristics and premature cardiovascular mortality in the study population for each National Health and Morbidity Survey	235
Table 4.18	Semi-parametric Cox Proportional Hazard regression model for premature cardiovascular disease mortality in Malaysia (N = 63,722)	239
Table 4.19	Hazard ratio of premature cardiovascular disease mortality for the interaction Cox models	246
Table 4.20	Parametric survival analysis models of premature cardiovascular disease mortality in Malaysia.....	249
Table 4.21	Sensitivity analysis treated of imbalance data for unadjusted survival analysis model of premature cardiovascular disease mortality in Malaysia	254
Table 4.22	Sensitivity analysis comparing treated imbalance data to original data for the adjusted premature cardiovascular disease mortality survival analysis model.....	256
Table 4.23	Summary of Fuzzy Delphi Method findings.....	279
Table 4.24	Suggested adjustment set for total effect of selected exposure toward premature CVD mortality	287
Table 4.25	DAG models for expert domain rate and testing model accuracy with available data.....	290
Table 4.26	Expert ranked the proposed DAG model	294

Table 4.27	Comparison model accuracy from data based on Logistic Regression models	297
------------	---	-----

LIST OF FIGURES

	Page
Figure 2.1 World Health Organization (WHO) framework of social determinants of health, 2010	36
Figure 2.2 Type of bibliometric analysis	40
Figure 2.3 Steps for systematic review study	44
Figure 2.4 Survival and Hazard Functions for selected distribution	54
Figure 2.5 Model classification in survival analysis	55
Figure 2.6 Terminology in directed acyclic graphs (DAGs)	60
Figure 2.7 Key concept of causal path in directed acyclic graphs (DAGs)	62
Figure 2.8 Framework of Causal Inference	64
Figure 3.1 Summary of methods for the overall study	68
Figure 3.2 Flow chart of the data cleaning process	115
Figure 3.3 Flow chart of study constructing DAG model for premature cardiovascular mortality	127
Figure 3.4 Summary of evidence synthesis for constructing DAG (ESC-DAG) protocol	129
Figure 4.1 Flow diagram of study selection process	138
Figure 4.2 Number of articles related to premature mortality	139
Figure 4.3 Collaboration of countries publishing articles related to premature mortality.	141
Figure 4.4 The top author's production over time.	144
Figure 4.5 Trending keywords related to premature mortality between 2002 and 2021	146
Figure 4.6 Thematic map of publications related to premature mortality	148
Figure 4.7 Co-occurrence keywords related to premature mortality	150

Figure 4.8	Flow diagram of the published articles evaluated for inclusion in this review	157
Figure 4.9	Distribution of selected studies by country	158
Figure 4.10	The YPLL rate of countries with available data	168
Figure 4.11	The SEYLL rate of countries with available data	171
Figure 4.12	Patterns of YPLL rate (a) and SEYLL rate (b) according to CVD types and WHO regions	173
Figure 4.13	Patterns of YPLL rate (a) and SEYLL rate (b) according to sex and WHO regions	174
Figure 4.14	Patterns of SEYLL rate by WHO regions.....	175
Figure 4.15	Flow diagram of the published articles evaluated for inclusion in this review	182
Figure 4.16	Forest plot of ASMR per 100,000 population from premature CVD mortality (all studies) with subgroup analysis by CVD types	188
Figure 4.17	Forest plot of ASMR per 100,000 population from premature CVD mortality (all studies) with subgroup analysis by sex (excluded both sex).....	190
Figure 4.18	Forest plot sex-specific premature mortality (ASMR per 100,000 population) from all CVD (ICD-10 codes I00-I99 or ICD-9 codes 350-459)	191
Figure 4.19	Forest plot sex-specific premature mortality (ASMR per 100,000 population) from IHD (ICD-10 codes I20-I25 or ICD-9 codes 410-414)	192
Figure 4.20	Forest plot sex-specific premature mortality (ASMR per 100,000 population) from cerebrovascular disease or stroke (ICD-10 codes I60-I69 or ICD-9 codes 430-438)	193
Figure 4.21	Forest plot of ASMR per 100,000 population from premature CVD mortality (all studies) with subgroup analysis by income country level.....	195

Figure 4.22	Forest plot of ASMR per 100,000 population from premature mortality among total CVD (ICD-10 codes I00-I99 or ICD-9 codes 350-459) with subgroup analysis by income country classification	196
Figure 4.23	Forest plot of ASMR per 100,000 population from premature CVD mortality (all studies) with subgroup analysis by study time.....	198
Figure 4.24	Forest plot time-stratified analysis of premature CVD mortality (ASMR per 100,000 population) among high income countries	199
Figure 4.25	Forest plot time-stratified analysis of premature CVD mortality (ASMR per 100,000 population) among middle (low and upper) income countries	200
Figure 4.26	Funnel plot to assess publication bias of overall studies (a) and total CVD (b).....	208
Figure 4.27	The ASMR trend from CVD in Malaysia from 2010 to 2021; (A) Overall population by age group; (B) Male by age group and (C) Female by age group	217
Figure 4.28	Joinpoint regression model of age standardized mortality rate (ASMR) per 100,000 from CVD in Malaysia, 2010-2021 for overall CVD (a) and by sex (b)	222
Figure 4.29	Joinpoint regression model of age standardized mortality rate (ASMR) per 100,000 from CVD in Malaysia, 2010-2021by age group in male (a) and female (b).....	224
Figure 4.30	Schoenfeld residual test and plot for each covariate before stratified Cox model.....	241
Figure 4.31	Schoenfeld residual test and plot for each covariate after stratified Cox model by age of entry	243
Figure 4.32	Parametric distribution fits to premature CVD mortality data by estimating baseline parametric models (without covariates)	248
Figure 4.33	Parametric distribution fittings for premature CVD mortality data after treatment using under-sampling method on imbalanced data .	253

Figure 4.34	Framework-related variables for premature cardiovascular mortality	269
Figure 4.35	ESC-DAG process (Mapping) - Implied graph (IG) for premature CVD from mapping stage	270
Figure 4.36	ESC-DAG process (Translation) - Translation of Implied graph to DAG for Premature CVD Mortality	273
Figure 4.37	ESC-DAG process (Integration) - The final I-DAG for premature cardiovascular mortality	275
Figure 4.38	DAG for premature CVD mortality from first round interview by expert domains	277
Figure 4.39	Final DAG for premature CVD mortality from evidence synthesis and expert consensus.....	285

LIST OF SYMBOLS

$\%$	Percentage
p	P value
$=$	Equal to
$>$	More than
$<$	Less than
\geq	More than and equal to
\leq	Less than and equal to
e	Exponential
β	Beta

LIST OF ABBREVIATIONS

AANS	American Association of Neurological Surgeons
AAPC	Average annual percentage change
ACLS	Aerobics Center Longitudinal Study
ACS	Acute coronary syndromes
AFT	Accelerated failure time
AIC	Akaike Information Criterion
AO	Abdominal obesity
APC	Annual percentage change
ASIR	Age-standardized incidence rate
ASMR	Age-standardized mortality rate
ASSIGN-SCORE	Assign Preventative Treatment - Systematic Coronary Risk Evaluation
ASYR	Age-standardized Years of Life Lost rates
AUC	Area under curve
BIC	Bayes Information Criterion
BMI	Body Mass Index
CDC	Centers for Disease Control
CENTRAL	Cochrane Central Register of Controlled Trials
CHD	Coronary heart disease
CI	Confidence intervals
COVID-19	Coronavirus disease 2019
CPI	Consumer Price Indices
CVA	Cerebral vascular accident
CVD	Cardiovascular disease
DAG	Directed Acyclic Graph
DALYs	Disability-adjusted life-years
DM	Diabetes mellitus
DOSM	Department of Statistics Malaysia
EPIC	European Prospective Investigation into Cancer and Nutrition
ESC-DAG	Evidence Synthesis for Constructing Directed Acyclic Graphs
ETR	Event Time Ratio
FDM	Fuzzy Delphi method
FRS	Framingham risk score
GBD	Global Burden of Disease
GO	General obesity
GPAQ	Global Physical Activity Questionnaire
HDL	High-density lipoprotein
HICs	High-income countries
HPL	Hyperlipidemia
HPT	Hypertension
HR	Hazard Ratio

ICD-10	International Statistical Classification of Diseases, 10th revision
I-DAG	Integrated Directed Acyclic Graph
IDE	Integrated Development Environment
IG	implant graph
IHD	Ischemic heart disease
IHME	Institute for Health Metrics and Evaluation
IPAQ	International Physical Activity Questionnaire
IQR	Interquartile range
JEPeM	Jawatankuasa Etika Penyelidikan (Manusia)
LICs	Low-income countries
LMICs	Low-middle-income countries
MET	Metabolic equivalent of task
MICs	Middle-income countries
MRFTT	Multiple Risk Factor Intervention Trial
NCD	Non-communicable disease
NCHS	National Center for Health Statistics
NHANES	National Health and Nutrition Examination Survey
NHMS	National Health and Morbidity Survey
NMRR	National Medical Research Register
NOS	Newcastle-Ottawa Scale
NRD	National Registration Department
NRIC	National Registration Identity Card
NSTEMI	Non-ST elevation (NSTEMI) myocardial infarction
OR	Odd Ratio
PAHO	Pan American Health Organization
PH	Proportional hazards
PHS	Physicians' Health Study
PI	Prediction interval
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROCAM	Prospective Cardiovascular Münster
PROSPERO	International Prospective Register of Systematic Reviews
PSC	Prospective Studies Collaboration
PURE	Prospective Urban Rural Epidemiology
QRISK	QRESEARCH cardiovascular risk
QUORUM	Quality of Reporting of Meta-Analyses
ROC	Receiver operating curve
SAS	Statistical Analysis Software
SCORE	Systematic Coronary Risk Evaluation
SCS	Seven Countries Study
SDG	Sustainable Development Goals
SDH	Social determinants of health
SE	Standard error
SES	Socioeconomic status
SEYLL	Standard Expected Years of Life Lost

SMR	Standardized mortality ratio
SPSS®	Statistical Package for the Social Sciences
STEMI	ST elevation myocardial infarction
TCS	Thai Cohort Study
TFN	Triangular Fuzzy Number
UK	United Kingdom
USA	United State of America
WC	Waist circumference
WHO	World Health Organization
WHO/ISH	World Health Organisation-International Society of Hypertension
WoS	Web of Science
YLL	Years of Life Lost
YPLL	Years of Potential Life Lost

LIST OF APPENDICES

Appendix A	PRISMA checklist (Study Phase 2)
Appendix B	Approval for the usage of mortality data from the Department of Statistics Malaysia (Study Phase 3)
Appendix C	Approval letter for the NHMS data from the Unit of Biostatistics & Data Repository, NIH (Study Phase 4)
Appendix D	Expert invitation letter for constructing a causal model for premature CVD mortality (Study Phase 5)
Appendix E	Questionnaire for Consensus-Building with Experts for Constructing a DAG Causal Pathway for Premature CVD Mortality (Study Phase 5)
Appendix F	Ethical approval from Medical Research & Ethics Committee, Ministry of Health Malaysia
Appendix G	Ethical approval from Jawatankuasa Etika Penyelidikan Manusia (JEPeM), USM
Appendix H	R code for statistical analysis
Appendix I	Excel worksheet template for calculating fuzzy Delphi method

**KEMATIAN PRAMATANG DARIPADA PENYAKIT
KARDIOVASKULAR: BEBAN GLOBAL DAN ANGGARAN KHUSUS
NEGARA MELALUI KAJIAN SEMAKAN, ANALISIS TREND,
PEMODELAN RAMALAN, DAN PENYEBAB KAJIAN INFERENS**

ABSTRAK

Penyakit kardiovaskular merupakan penyebab utama kematian pramatang di seluruh dunia, dengan beban penyakitnya kian meningkat dalam beberapa tahun terkini. Walaupun banyak penyelidikan telah dilakukan berkaitan dengan penyakit kardiovaskular, namun, terdapat jurang dalam mengenalpasti perkaitan kematian pramatang dengan kardiovaskular. Sebagai contoh, terdapat maklumat yang tidak mencukupi mengenai aktiviti penyelidikan global dan anggaran kumpulan beban kematian pramatang akibat penyakit kardiovaskular dari sumber kajian yang sedia ada. Selain itu, anggaran beban ini dalam negara dan profil risiko penyakit kardiovaskular yang khusus untuk setiap negara turut kurang. Justeru itu, tesis ini bertujuan untuk menganalisis dan memahami secara menyeluruh pelbagai aspek kematian pramatang berkaitan penyakit kardiovaskular, termasuk aktiviti penyelidikan secara global, membuat anggaran global beban kematian pramatang disebabkan penyakit kardiovaskular, mengenalpasti trend dan profil risiko bagi kematian pramatang penyakit kardiovaskular secara spesifik di dalam negara, serta membina model laluan kausa bagi kematian pramatang penyakit kardiovaskular. Tesis ini terdiri daripada lima kajian yang berkaitan bertujuan untuk menjawab semua objektif penyelidikan. Kajian pertama meringkaskan maklumat penerbitan global yang berkaitan dengan kematian pramatang melalui analisis bibliometrik. Analisis bibliometrik menunjukkan bahawa penyelidikan kematian pramatang telah meningkat secara ketara dalam sedekad yang

lalu; namun, terdapat perbezaan dalam hasil penyelidikan antara negara berpendapatan tinggi (HICs) dan negara berpendapatan rendah dan pertengahan (LMICs). Kajian ini menekankan keperluan untuk usaha kerjasama di negara LMICs. Dalam kajian kedua, terdiri daripada dua bahagian: kajian 2a merupakan kajian ulasan sistematik yang menumpukan kepada kematian pramatang penyakit kardiovaskular menggunakan indikator jumlah tahun hayat (YLL), diikuti oleh kajian 2b, meta-analisis berdasarkan indikator kadar kematian diselaraskan mengikut umur (ASMR). Hasil kajian 2a dan 2b ini menunjukkan beban yang berterusan kematian pramatang penyakit kardiovaskular, terutamanya di negara LMICs dan di kalangan lelaki. Kajian ketiga melibatkan analisis trend menggunakan data kematian spesifik negara dari Jabatan Perangkaan Malaysia dan menggunakan kaedah model regresi joinpoint. Kajian ini mendapati peningkatan kadar kematian pramatang penyakit kardiovaskular di Malaysia secara berterusan dalam satu dekad yang lalu, menunjukkan beban yang berterusan di negara ini. Dalam kajian keempat, satu kajian pemodelan dijalankan untuk menilai profil risiko kematian pramatang penyakit kardiovaskular dengan menggunakan pelbagai model survival. Kajian ini menggunakan data daripada Tinjauan Kebangsaan Kesihatan dan Morbiditi (NHMS), Malaysia yang dijalankan pada tahun 2006, 2011, dan 2015, dan dipadankan dengan rekod kematian sehingga tahun 2021. Model survival semi-parametrik dan parametrik kedua-duanya menunjukkan status sosioekonomi dan faktor risiko yang boleh diubah merupakan faktor risiko penting terhadap kematian pramatang penyakit kardiovaskular di Malaysia. Kajian ini menekankan keperluan untuk intervensi berpusat bagi mengurangkan kesan ini secara berkesan. Akhirnya, kajian kelima melibatkan satu kajian kesimpulan kausa dimana fasa ini menunjukkan proses secara sistematik untuk membina model Directed Acyclic Graph (DAG) bagi kematian pramatang penyakit kardiovaskular. Kajian ini memperkenalkan pendekatan baru yang

mengintegrasikan bukti sintesis dan pendapat pakar menggunakan kaedah Fuzzy Delphi untuk membina model DAG, diikuti dengan menguji model dengan data yang ada. Proses langkah demi langkah yang ditunjukkan untuk membina model kausa ini, menawarkan pendekatan sistematik, praktikal, transparan, dan menyumbang kepada metodologi DAG. Secara keseluruhan, melalui metodologi yang teliti, dapatan tesis ini menyumbang kepada perkembangan karya sastera tentang kematian pramatang dari penyakit kardiovaskular dan memberikan input dan kerangka nilai yang berharga untuk memberi informasi kepada strategi kesihatan awam dan membimbing usaha penyelidikan masa depan dalam menangani kematian pramatang penyakit kardiovaskular.

**PREMATURE MORTALITY FROM CARDIOVASCULAR DISEASE:
GLOBAL BURDEN AND COUNTRY-SPECIFIC ESTIMATE THROUGH
REVIEW STUDY, TREND ANALYSIS, PREDICTION MODELLING, AND
CAUSAL INFERENCE STUDY**

ABSTRACT

Cardiovascular disease (CVD) is a leading cause of premature mortality worldwide, with its burden growing in recent years. Despite considerable research into CVD, there are noticeable gaps in addressing premature mortality associated with it. For example, there is insufficient information on global research activity and pool estimation of the global burden of premature CVD mortality from existing literature. Additionally, within-country estimates of this burden and country-specific CVD risk profiles are lacking. Therefore, this thesis aims to comprehensively analyze and understand various aspects of premature CVD mortality, including global research activities, estimating the global burden of CVD related to premature mortality, examining country-specific trends and risk profiles for premature CVD mortality, and constructing causal pathways for premature CVD mortality. This thesis consists of five interrelated studies aimed at addressing all research objectives. The first study summarizes the global publication information related to premature mortality through bibliometric analysis. The bibliometric analysis shows that premature mortality research has increased substantially in the past decade; however, there have been disparities in research output between high-income countries (HICs) and low-middle-income countries (LMICs), emphasizing the need for collaborative efforts in LMICs. In the second study, consisting of two parts: Study 2a, a systematic review concentrating on premature CVD mortality using the Years of Life Lost (YLL)

indicator, followed by study 2b, a subsequent meta-analysis based on the age-standardized mortality rate (ASMR) indicator. Studies 2a and 2b underscore the persistent burden of premature CVD mortality, particularly in LMICs and among men, emphasizing the importance of targeted interventions and public health strategies. The third study involves trend analysis utilizing Malaysia country-specific mortality data from Department of Statistics Malaysia employing a joinpoint regression model. This study reveals a persistent increasing trend of premature CVD mortality rates in Malaysia over the past decades, underscoring the ongoing burden in the country. In study four, a modeling study is conducted to assess the premature CVD risk profile by applying various survival models. This study utilizes data from Malaysia's National Health and Morbidity Survey (NHMS) conducted in 2006, 2011, and 2015, linked with mortality records through 2021. The semi-parametric and parametric survival models both highlight the considerable impact of socioeconomic status and modifiable risk factors on premature CVD mortality in Malaysia. Finally, the fifth study entails a causal inference study systematically demonstrating the construction of a Directed Acyclic Graph (DAG) model for premature CVD mortality. This study introduces a new approach integrating evidence synthesis and expert opinion using the Fuzzy Delphi method to construct the DAG model, followed by testing the model with available data. This step demonstrates a step-by-step process for constructing a causal model, offering a systematic, practical, and transparent approach, contributing to DAG methodology. In summary, through meticulous methodologies, the findings of this thesis contribute to the growing body of literature on premature mortality from CVD and provide valuable insights and frameworks for informing public health strategies and guiding future research endeavours in tackling premature CVD mortality.

CHAPTER 1

INTRODUCTION

1.1 Background of the study

Cardiovascular diseases (CVDs) encompass a wide range of disorders, from conditions affecting the heart muscle to those impacting the blood vessels supplying the heart, brain, and other vital organs (Gaziano *et al.*, 2006). Globally, CVD continues to be the leading cause of morbidity and mortality, imposing significant social and economic burdens. In 2017, an estimated 17.8 million people died from CVDs worldwide, resulting in the loss of 330 million years of life and an additional 35.6 million years lived with disability (Kyu *et al.*, 2018; Roth *et al.*, 2018a). Notably, approximately 80% of CVD deaths occur in low- and middle-income countries (LMICs) (WHO, 2021a).

In addition to the staggering number of deaths caused by CVDs, a significant portion comprises premature mortality. Premature mortality is defined as a death that occurs before the average age of death in a particular population (National Cancer Institute, 2021a). It is a measure of unfulfilled life expectancy. The trends in premature mortality globally from 1990–2017 have shown a 41% decrease in communicable diseases and neonatal disorders but a 40% increase in non-communicable diseases (NCDs) (IHME., 2017).

According to the Global Burden of Disease (GBD) study, CVDs, the main types of NCDs, continue to dominate as the leading causes of premature mortality worldwide (Roth, Huffman, *et al.*, 2015a). Principally, ischemic heart disease (IHD) and cerebrovascular disease account for approximately one-third of all global premature deaths (Roth, Huffman, *et al.*, 2015b; Vos *et al.*, 2020). Together with

cancer, respiratory diseases, and diabetes, they collectively contribute to over 80% of all premature mortality due to NCDs (WHO, 2022). LMICs bear a disproportionately high burden of premature mortality compared to high-income countries (HICs) (WHO, 2022). Furthermore, CVDs in LMICs are significant contributors to the escalating burden of NCDs, accounting for approximately 86% of all premature NCD deaths globally (WHO, 2022).

Similarly, in Malaysia, CVD has been the leading cause of death since the 1990s and remains a significant threat, contributing to around 35% of premature mortality in the country (WHO, 2018a; Noor, Muzafar & Khalidi, 2020). In 2018, IHD continued as the predominant cause of premature mortality among Malaysians, while stroke was positioned as the fourth ranking (Khaw *et al.*, 2023).

Premature CVD mortality is influenced by various risk factors, categorized as non-modifiable and modifiable. Non-modifiable factors like age, sex, and ethnicity are beyond our control but affect CVD susceptibility (Choudhury *et al.*, 2015; Lotufo, 2015; Mohammadnezhad *et al.*, 2016; J. Zhang, Jin, Jia, Li & Z.-J. Zheng, 2021). On the other hand, modifiable risk factors allow targeted interventions and lifestyle changes to reduce CVD morbidity and mortality.

Among the key modifiable risk factors, metabolic factors such as diabetes, hypertension, hyperlipidemia, and obesity significantly contribute to premature CVD mortality (Eslami *et al.*, 2019; Yusuf *et al.*, 2020; Joseph *et al.*, 2022). Furthermore, beyond metabolic factors, various behavioral risk factors significantly escalate the incidence of premature CVD mortality. These include smoking (Eslami *et al.*, 2019; Joseph *et al.*, 2022), excessive alcohol consumption (Leon *et al.*, 2007), physical inactivity, and sedentary lifestyles, all of which play critical roles in elevating CVD

mortality rates (WHO, 2021c). Effective management and early detection of individuals with CVD and its risk factors, particularly behavioral and modifiable ones, necessitate targeted interventions to mitigate premature mortality.

1.2 Problem statement

Although extensive research has been conducted on CVD from various angles, there remains a need for more studies focusing on CVD-related premature mortality both globally and within specific countries. From a global perspective, there is insufficient research on the overall global research activity concerning premature mortality and the estimation of the global burden specifically attributable to CVD-related premature deaths based on existing literature.

Firstly, while the academic literature on premature mortality has expanded recently, there is a lack of comprehensive information regarding the total number and distribution of published articles in this area. Bibliometric analysis, a valuable method for exploring the distribution of publications, evaluating research progress, uncovering emerging research trends, and identifying relevant topics related to any field of study, could provide valuable insights into the current state of research activity. Extracting knowledge from existing studies is crucial for identifying current developments and future trends, which is essential for assessing the relevance of topics related to premature mortality.

Secondly, despite the extensive investigation of CVD mortality through systematic reviews and meta-analyses, premature mortality has frequently been overlooked as an outcome measure. These analyses typically focus on establishing predictors of increased mortality, relative risks, or risk factors for cause-specific mortality (Aviña-Zubieta *et al.*, 2008; Nocon *et al.*, 2008; Huang *et al.*, 2014; Qiu *et*

al., 2021) rather than emphasizing premature death age-related to CVD. Meanwhile, the GBD study, the most widely used global estimate of premature mortality, has derived estimates of global premature CVD mortality with missing or limited-quality mortality data in some countries, particularly in the poorest regions (Roth *et al.*, 2018b; Stanaway *et al.*, 2018). To the best of researcher knowledge, there has yet to be a comprehensive systematic review and meta-analysis that synthesizes the available evidence on premature CVD mortality, specifically focusing on estimating a pooled effect size using metrics such as years of life lost (YLL) and age-standardized mortality rate (ASMR), common methods for measuring the burden of premature mortality. Incorporating diverse parameters in such an analysis would offer a thorough evaluation of premature CVD mortality, thus contributing significantly to the development of global strategies aimed at addressing this critical public health concern.

Turning to a country-specific perspective, investigating trends in CVD mortality within countries, especially in LMICs like Malaysia, is crucial. However, Malaysia's CVD mortality trends have remained unexplored over the past decade. Analyzing Malaysia's data from vital statistics would provide valuable insights for policy and intervention strategies. While joinpoint analysis is a statistical method commonly used to detect changes in trends or patterns in data over time, including the impact of CVD mortality rates, it has not been applied to examine the trend and any change in the mortality rate from CVD in Malaysia over the past decade.

Furthermore, the risk factors for premature CVD mortality vary across regions and income levels. The PURE project, studying 155,722 participants across 21 countries, found differing impacts based on income levels (Yusuf *et al.*, 2020). In

middle-income countries (MICs), hypertension, metabolic, and behavioral factors are significant, with low education having a pronounced effect compared to HICs. In low-income countries (LICs), metabolic factors, household air pollution, and poor diet are key risks (Yusuf *et al.*, 2020). In South Asia, a cohort study of 33,583 individuals highlighted the impact of low education, weak physical strength, and poor diet quality (Joseph *et al.*, 2022). Similarly, in China, hypertension remains a primary risk factor, with low education ranking high for both CVD and mortality (Li *et al.*, 2022). Such diversity in risk profiles likely contributes to variations in CVD mortality rates across the Asian region. Malaysia's diverse population encompasses various ethnicities, each with distinct genetic predispositions and lifestyle habits that can influence CVD risk profiles. Despite existing research on CVD risk factors, the study of premature CVD mortality in Malaysia remains limited, especially in employing survival analysis to model modifiable risk factors associated with premature CVD mortality using population-based data. Such information is crucial for informing policy formulation and public health strategies for controlling and preventing CVDs in Malaysia.

Lastly, while numerous models predict CVD mortalities (Lynch *et al.*, 1996; Eslami *et al.*, 2017; Joseph *et al.*, 2022), most risk prediction and causation models in epidemiology rely on additive combinations of risk factors within a regression model framework. However, these models may not accurately represent the causal pathways underlying disease development. It's crucial in epidemiological studies to distinguish between prediction and causal approaches to ensure valid interpretation of the evidence.

Causal inference approaches explicitly consider causality and describe the influence of variables on disease outcomes by accounting for causal pathways between

them (Pearl, 2009). Directed Acyclic Graphs (DAGs) serve as a valuable tool to represent causal models (Pearl & others, 2000; Pearl, Glymour & Jewell, 2016). A DAG shows causal links between variables using arrows connecting nodes. Theory and background knowledge are central to DAG construction, and drawing a DAG before analysis is crucial for guiding the selection of confounding variables in a study (Robins & Wasserman, 1999; Robins, 2001; Hernán *et al.*, 2002; Sauer *et al.*, 2013).

Despite the importance of DAGs, there is currently no systematic documentation of standard methods or protocols for constructing them, particularly in epidemiological studies. Although Ferguson *et al.* (2020) introduced a systematic approach called Evidence Synthesis for Constructing Directed Acyclic Graphs (ESC-DAG), there remains a need for practical guidance on drawing DAGs, incorporating the experience and knowledge of domain experts (Ferguson *et al.*, 2020). Therefore, a comprehensive methodology that integrates evidence synthesis and expert domains is necessary to provide practical guidance on drawing DAGs and applying them effectively in data analysis. Researchers can use this approach as an additional tool for selecting adjustment variables when analyzing epidemiological data.

Therefore, this thesis aims to address existing gaps in the literature through meticulous methodological analyses and study design to estimate the global and country-specific burdens, as well as the risk factors, for premature CVD mortality. From a global perspective, this study adopts a comprehensive approach, utilizing bibliometric analysis, systematic review, and meta-analysis to investigate global research activity on premature mortality and estimate the global burden of CVD-related premature mortality. For the country-specific perspective, Malaysia's vital registration data is used in a modeling study employing a joinpoint regression model

to analyze trends in CVD mortality rates in Malaysia. To estimate specific risk factors and construct a causal model, survival analysis—incorporating both semi-parametric and parametric survival models—is applied to assess the CVD risk profile using population-based surveys from Malaysia linked with mortality data. Finally, this thesis employs causal inference methodology, demonstrating a systematic process for constructing a DAG as a causal model for premature CVD mortality. This integrates both evidence synthesis and expert consensus to provide a robust framework for understanding causal pathways.

1.3 Research questions

The primary research question for this study is: What are the global estimates of premature CVD mortality, and what are the country-specific trends, risk factors, and causal pathways? This overarching question is explored through the following specific research questions:

- I. What are the characteristics of publications on premature mortality, particularly concerning the distribution of publications, citations, countries involved, collaboration patterns, author productivity, trending keywords, and relevant research topics?
- II. What are the characteristics and findings of studies investigating premature CVD mortality, explicitly focusing on years of life lost?
- III. What is the pooled estimate of age-standardized mortality rates for premature CVD mortality derived from a systematic review and meta-analysis of relevant studies?
- IV. What are the temporal trends in age-standardized mortality rates related to CVD in Malaysia between 2010 and 2021?

- V. What prognostic factors are associated with premature CVD mortality in Malaysia, specifically focusing on modifiable risk factors?
- VI. What causal pathways underlie premature CVD mortality, and how can they be constructed through a comprehensive approach integrating evidence synthesis and expert domain knowledge utilizing the DAGs method?

1.4 Research objectives

1.4.1 General objective

The general objective of this study is to determine the patterns and burdens of premature mortality from CVD, as well as its contributing factors and causal pathways.

1.4.2 Specific objectives

- I. To explore the characteristics of the publications on premature mortality from all causes of death, in terms of the number of publications, citations, countries, collaboration, and author's productivity and further identify the trending keyword and relevant research topic.
- II. To identify the studies on premature CVD mortality and synthesise their findings on years of life lost based on the regional area, main CVD types, sex, and study time.
- III. To estimate the pooled age-standardized mortality rates of premature CVD mortality through a comprehensive systematic review with meta-analysis, providing both global estimates and pooled estimates by sex, major types of CVD, income, country level, and study period.
- IV. To explore the trend and changes in CVD mortality rates in Malaysia using age-standardized mortality rates from 2010 to 2021 among all age groups, including premature mortality, using Malaysia's vital registration data.

- V. To determine the prognostic factors associated with premature CVD mortality in Malaysia, with a particular focus on modifiable risk factors among Malaysian adults aged below 70 years who died from CVD, using Malaysia's population-based data linked with death records.
- VI. To construct causal pathways for premature CVD mortality based on a comprehensive approach integrating evidence synthesis and expert domain knowledge through DAGs.

1.5 Overview of studies

The PhD project comprises of five interrelated studies aimed at addressing all research objectives. The first study focused on conducting a bibliometric analysis of premature mortality to enhance understanding of global research activities on premature mortality across all causes of death. In the second study, a systematic review was conducted specifically on premature mortality due to CVD, consisting of two parts: a systematic review concentrating on premature CVD mortality using the YLL indicator, followed by a subsequent meta-analysis based on the ASMR indicator.

The third study involved trend analysis utilizing country-specific mortality data from vital registration employing a joinpoint regression model. In the fourth study, a modeling analysis was conducted to assess the premature CVD risk profile by applying various survival models. Finally, the fifth study entailed a causal inference study systematically demonstrating the construction of a DAG model for premature CVD mortality. An overview of each study, including methods and main findings, is provided below.

1.5.1 Study 1: Bibliometric analysis on premature CVD mortality

This study aimed to explore the characteristics of publications on premature mortality, including publication numbers, citations, countries involved, collaboration patterns, author productivity, and identifying trending keywords and relevant research topics. Articles related to premature mortality data were retrieved from the Web of Science (WoS) database using specific search terms. A Bibliometrix package from R software was used to perform bibliometric analyses. A total of 1,060 original research articles and reviews were analyzed, revealing a significant increase in publications over the past decade, with the United States of America (USA) leading in research output and collaboration. Thematic mapping highlighted CVD as the primary research domain, with potential areas for future research including air pollution and neurodegeneration. These findings provide valuable insights for scholars, potential collaborators, and future research endeavors in understanding and addressing premature mortality globally.

1.5.2 Study 2a: Systematic review on years of life lost from CVD mortality

The objective of this study was to identify studies on premature CVD mortality and synthesize their findings using the YLL indicator, focusing on regional areas, major CVD types, sex, and study period. A systematic review was conducted across multiple databases including PubMed, Scopus, WoS, and the Cochrane Central Register of Controlled Trials (CENTRAL), yielding 40 eligible studies. These studies, representing various WHO regions, reported both Years of Potential Life Lost (YPLL) and Standard Expected Years of Life Lost (SEYLL), methods that are commonly used to calculate YLL. Results showed varying rates across income countries, with middle income countries (MICs) experiencing higher rates. Over three decades, there was a slight increase in YPLL and SEYLL rates for overall CVD and IHD, while

cerebrovascular disease rates slightly decreased. Males exhibited higher rates compared to females, particularly in the increase over time. The study concludes that premature CVD mortality remains a significant burden for MICs, necessitating urgent strategies to mitigate this mortality gap.

1.5.3 Study 2b: Meta analysis of ASMR from premature CVD Mortality

In this study, the pooled ASMR of premature CVD mortality was estimate. Through a systematic review of articles published up to October 2022 in PubMed, Scopus, WoS, and CENTRAL databases, 15 studies meeting the inclusion criteria were identified. The pooled estimate of ASMR was calculated using random-effects meta-analysis in R software. The global ASMR for premature CVD mortality was estimated at 96.04 per 100,000 people, with subgroup analyses revealing higher rates for IHD compared to cerebrovascular disease or stroke. Sex-specific differences showed higher ASMRs for males than females, while MICs exhibited significantly higher ASMRs compared to HICs. Stratifying by age group indicated that the age groups of 20–64 years and 30-74 years had a higher ASMR than the age group of 0-74 years. The multivariable meta-regression model suggested significant differences in the adjusted ASMR estimates for all covariates except study time. The findings underscore the persistent burden of premature CVD mortality, particularly in MICs, emphasizing the need for targeted interventions in these populations.

1.5.4 Study 3: Trend of ASMR from CVD in Malaysia

This study aimed to investigate trends and changes in CVD mortality rates in Malaysia from 2010 to 2021. Data on the Malaysian population and mortality were obtained from the Department of Statistics Malaysia (DOSM). ASMRs for CVD per 100,000 population were calculated using the World Health Organization (2000-2025) standard population using the direct method. ASMRs were stratified by sex, age groups

(including premature mortality ages 30-69 years), and CVD types. The annual percent change (APC) and average annual percent change (AAPC) of ASMRs were estimated using an advanced regression modelling known as the joinpoint regression model. Results indicated a notable increase in all age CVD ASMRs from 93.1 to 147.0 per 100,000 population during the study period, with an AAPC of 3.6%. The increase was particularly prominent between 2015 and 2018. The study highlights the persistent burden of premature CVD mortality (age 30-69) in Malaysia, particularly among males, underscoring the ongoing challenge of premature CVD mortality in the country.

1.5.5 Study 4: Prognostic Factors for Premature CVD Mortality in Malaysia

This study employs advanced semi and full parametric survival analysis to model modifiable risk factors among Malaysian adults. Utilizing data from Malaysia's National Health and Morbidity Survey (NHMS) conducted between 2006 and 2015, linked with mortality records, the study analyzes a cohort of 63,722 individuals aged 18 to 69. Among them, 1.4% experienced premature CVD mortality, with an age-standardized incidence rate (ASIR) of 1.80 per 1000 person-years. Employing six survival models, including a semi-parametric Cox proportional hazard model and five parametric survival models (Exponential, Weibull, Gompertz, log-normal and log-logistic), the study found that the best-fit survival models were the stratified Cox model by age and the log-normal accelerated failure time model. Males, rural residents, and those with lower education levels faced higher risks, while diabetes, hypertension, hypercholesterolemia, smoking, and abdominal obesity were associated with increased mortality risk and reduced survival time. These findings underscore the significant impact of socioeconomic status and modifiable risk factors on premature CVD mortality, highlighting the need for targeted interventions.

1.5.6 Study 5: Constructing DAG model for Premature CVD Mortality

This study presents a novel methodology for constructing a causal model for premature CVD mortality, integrating evidence synthesis and expert consensus across three stages. In Stage 1, the research team develops a draft DAG model based on the ESC-DAG framework. Subsequently, in Stage 2, the model undergoes validation and consensus-building with 12 experts through two rounds of interviews, using the Fuzzy Delphi method (FDM) for final consensus. In this phase, the initial draft DAG model was revised based on expert consensus, and the suggested minimal adjustment set from this final DAG model guided the selection of confounding variables for adjustment as a general reference for epidemiological studies on premature CVD mortality. For example, to estimate the total effect of diabetes on premature CVD mortality, the suggested minimal adjustment set included age, dietary pattern, genetic or family history, sex hormones, and physical activity. In the final stage, the model is applied to guide data analysis, proposing three feasible DAG models using population-based survey data linked with death records. Testing different DAG models in this stage revealed consistent agreement between expert ratings and data accuracy from regression models for the suggested adjustment set. The study concludes that integrating evidence synthesis with expert opinion enhances the construction of DAGs, providing a more robust and realistic representation of the causal model and aiding in identifying appropriate adjustment strategies for epidemiological analysis.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter provides a comprehensive review of cardiovascular diseases and related topics, including methodologies used in this study. The chapter begins with an overview of the definition, classification, and epidemiology of CVDs, followed by an exploration of premature mortality, including its measurement methods and the global and country-specific burdens associated with premature CVD mortality. The chapter then delves into the existing literature on factors associated with CVD mortality. Subsequently, methodologies employed in review studies, such as bibliometric analysis and meta-analysis, are discussed. This is followed by a review of trend analysis techniques and prediction modeling methodologies, with a specific focus on survival modeling. Finally, the chapter provides an overview of causal inference frameworks, specifically addressing the DAG method.

2.2 Cardiovascular diseases

2.2.1 Definition CVD

Cardiovascular diseases or CVD encompasses a broad spectrum of disorders, ranging from conditions affecting the cardiac muscle to those impacting the vascular system that supplies blood to essential organs such as the heart, brain, and other vital organs (Gaziano *et al.*, 2006). In general terms, CVD refers to conditions affecting the heart or blood vessels, often associated with the accumulation of fatty deposits within the arteries (atherosclerosis) and an increased susceptibility to blood clots (WHO, 2021a). CVD stands as one of the four primary non-communicable diseases (or NCDs),

alongside cancers, respiratory diseases, and diabetes (WHO, 2022). NCDs are characterized by conditions that usually progress gradually over time and are not primarily caused by acute infections. Unlike communicable diseases, NCDs are not spread through infection or direct contact with other individuals but are often linked to unhealthy behaviors (PAHO., 2023). These diseases frequently lead to long-term health complications, necessitating ongoing treatment and care. They pose a significant threat to health and development, with NCDs claiming the lives of 41 million people annually, representing 74% of all deaths globally (WHO, 2022).

2.2.2 Main type of CVD

There are several types of CVDs, including: 1) coronary heart disease, which affects the vessels supplying the heart muscle; 2) cerebrovascular disease, impacting vessels supplying the brain; 3) peripheral arterial disease, affecting vessels supplying the arms and legs; 4) rheumatic heart disease, resulting from damage to the heart muscle and valves; 5) congenital heart disease, involving birth defects that affect the heart's development and function; and 6) deep vein thrombosis and pulmonary embolism, entailing blood clots in leg veins that may dislodge and migrate to the heart and lungs (WHO, 2021a). The most common causes of CVD morbidity and mortality are ischemic heart disease and cerebrovascular diseases, or stroke (WHO, 2021b). This review will describe these two main types of CVD.

2.2.2(a) Ischemic heart disease

The term ischemic heart disease or IHD falls within the broader category of coronary heart disease. "Ischemic" denotes a condition where an organ, in this case, the heart, does not receive adequate blood and oxygen supply. Ischemic heart disease specifically refers to heart issues stemming from narrowed coronary arteries, which are responsible for supplying blood to the heart muscle (Institute of Medicine., 2010). It

encompasses a range of clinical syndromes characterized by myocardial ischemia, which signifies an imbalance between the supply of blood to the myocardium (heart muscle) and its demand for oxygen (Steenbergen & Frangogiannis, 2012).

The presentations of IHD vary depending on the duration, severity, and acuteness of the ischemic episodes. Acute coronary syndromes (ACS), a range of clinical conditions including unstable angina, non-ST-elevation myocardial infarction, and ST elevation myocardial infarction, stem from a sudden critical reduction in coronary blood flow. Conversely, in chronic IHD, the presence of coronary lesions that limit blood flow impedes the heart's ability to meet increased myocardial oxygen demands, leading to the development of angina pectoris—a transient discomfort felt in the chest and nearby areas (Steenbergen & Frangogiannis, 2012). The primary two manifestations of IHD are angina and acute myocardial infarction (Gaziano *et al.*, 2006).

2.2.2(b) Stroke

The term "cerebrovascular" consists of two components: "cerebro," pertaining to the large part of the brain, and "vascular," relating to arteries and veins. Together, "cerebrovascular" refers to the blood flow within the brain (AANS., 2024). A stroke, a subtype of cerebrovascular disease also known as a cerebrovascular accident, is an acute medical emergency characterized by a sudden compromise in cerebral perfusion or vasculature (Khaku & Tadi, 2024). It occurs when blood flow to certain areas of the brain is obstructed (ischemic stroke) or when a blood vessel ruptures (hemorrhagic stroke), resulting in damage to brain tissue.

Ischemic strokes, the most common type, account for about 85% of cases (Mozaffarian *et al.*, 2016). Ischemic strokes are classified as thrombotic or embolic. Thrombotic strokes happen when a blood clot obstructs a brain artery, while embolic

strokes occur when plaque or a clot dislodges and blocks a downstream artery (AANS., 2024). Hemorrhagic strokes result from weakened vessel rupture due to factors like high blood pressure, aneurysm rupture, or medication side effects. Intracerebral hemorrhage involves bleeding directly into brain tissue, often forming a clot, while subarachnoid hemorrhage fills the cerebrospinal fluid spaces around the brain, both posing serious risks (AANS., 2024).

Stroke can manifest with various symptoms, including sudden weakness or paralysis on one side of the body, difficulty speaking or understanding speech, and vision problems (Abiodun, 2018). Immediate medical attention is crucial to minimize brain damage and improve outcomes. Treatment options may include clot-busting medications, surgical interventions, and rehabilitation therapies. Prevention strategies include managing risk factors such as high blood pressure, diabetes, and smoking, as well as adopting a healthy lifestyle with regular exercise and a balanced diet (CDC., 2022).

2.2.3 Classification of CVD (codes)

Established by the World Health Organization, the International Statistical Classification of Diseases, 10th revision (ICD-10) serves as the global standard for recording and analyzing information on both mortality and morbidity (illness) statistics across various diseases (National Center for Health Statistics, 2022). ICD-10 can be used to code and classify morbidity data from the inpatient and outpatient records, physician offices, and most National Center for Health Statistics (NCHS) surveys. This standardized system facilitates the efficient storage, retrieval, analysis, and interpretation of health data, enabling meaningful comparisons between different countries. This allows for a comprehensive understanding of global health trends and facilitates informed decision-making in healthcare policies and interventions.

The ICD-10 code range for CVD is I00-I99, encompassing a diverse range of conditions affecting the heart and circulatory system. This includes specific codes for main CVD types: ICD-10 codes I20-I25 for Ischemic heart diseases (e.g., coronary heart disease) and codes I60-I69 for Cerebrovascular diseases (e.g., stroke). For further details on specific codes, refer to **Table 2.1**.

Table 2.1 The International Classification of Diseases, 10th revision (ICD-10) for main CVD

CVD types	Codes	Disease
Ischemic heart diseases	<u>I20</u>	Angina pectoris
	<u>I21</u>	Acute myocardial infarction
	<u>I22</u>	Subsequent ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction
	<u>I23</u>	Certain current complications following ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction (within the 28-day period)
	<u>I24</u>	Other acute ischemic heart diseases
Heart failure	<u>I25</u>	Chronic ischemic heart disease
Cerebrovascular diseases / Stroke	I50	Heart failure
	<u>I60</u>	Nontraumatic subarachnoid hemorrhage
	I61	Nontraumatic intracerebral hemorrhage
	I62	Other and unspecified nontraumatic intracranial hemorrhage
	<u>I63</u>	Cerebral infarction/CVA/stroke
	<u>I65</u>	Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction
	<u>I66</u>	Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction
	<u>I67</u>	Other cerebrovascular diseases
	<u>I68</u>	Cerebrovascular disorders in diseases classified elsewhere
	<u>I69</u>	Sequelae of cerebrovascular disease

Source: (National Center for Health Statistics, 2022)

2.2.4 Epidemiology of CVD

2.2.4(a) Incidence, burden, and trend of CVD

Cardiovascular diseases primarily IHD and stroke, are the world's leading cause of death, disability, and premature mortality, as well as a major contributor to rising healthcare costs (Roth *et al.*, 2020; Vos *et al.*, 2020). This burden falls disproportionately on LMICs, where CVD accounts for nearly 80% of global deaths (WHO, 2021a). Since 1990, CVD cases have more than doubled, reaching 523 million in 2019. Similarly, CVD deaths have nearly doubled, reaching 17.8 million in 2017, and years of life lost to disability have significantly increased (Roth *et al.*, 2020).

2.2.4(b) The global burden of IHD

Ischemic heart disease remains a major public health challenge worldwide, claiming over 7 million lives in 2010 alone (American Heart Association., 2020). While improved survival rates for acute myocardial infarction (AMI) have emerged, many individuals live with chronic symptoms associated with stable angina pectoris or ischemic heart failure. Notably, the 2010 Global Burden of Diseases, Injuries, and Risk Factors study employed disability-adjusted life-years (DALYs) to quantify the burden of IHD and other prevalent diseases (Moran *et al.*, 2014). Their findings revealed a complex interplay: despite declining age-adjusted IHD incidence and mortality in most regions, population growth and aging have led to an overall increase in the global burden of IHD since 1990. Interestingly, the study also highlighted that the global IHD burden is primarily driven by fatal IHD with non-fatal IHD contributing less significantly, and acute non-fatal IHD (AMI or unstable angina) having the least impact (Moran *et al.*, 2014).

2.2.4(c) The global burden of stroke

Stroke is the second leading cause of death and the fourth leading cause of disability globally, measured in disability-adjusted life years (WHO, 2008). In 2005, the WHO estimated 5.7 million deaths and 16 million new strokes, leaving an estimated 62 million stroke survivors (Strong, Mathers & Bonita, 2007). By 2030, these numbers are projected to rise to 7.8 million deaths and 23 million new strokes, respectively (WHO, 2008). Though stroke remains a major cause of death in developed countries, LMICs now experience higher rates of first-time strokes compared to HICs, with this trend beginning around 2000 (Feigin *et al.*, 2009).

2.2.4(d) Epidemiological transitions

The global burden of CVD and its risk factors continues to increase due to ongoing epidemiological transitions (Yusuf *et al.*, 2001). Prior to 1900, infectious diseases and malnutrition were leading causes of death. CVD was not as prevalent, with rheumatic heart disease being the main concern, often stemming from infections or malnutrition. As infectious diseases lessened and nutrition improved, a second stage emerged with hypertensive illnesses like hemorrhagic stroke and hypertensive heart disease gaining prominence. Now, in the third stage, longer life expectancy, high-fat diets, smoking, and sedentary lifestyles drive the rise of NCDs. IHD and stroke now dominate mortality rates, especially among those under 50 (Yusuf *et al.*, 2001).

2.3 Premature mortality

2.3.1 Definition

Premature mortality, defined as death occurring before the average age of death within a specific population (National Cancer Institute, 2021b), serves as a pivotal indicator of population health. It reflects the unfulfilled life expectancy, which estimates

the additional years an individual of a certain age can typically anticipate living (Bezy, 2024). However, life expectancy varies significantly across regions, influenced by factors such as sex, age, race, and geographic location. Notably, life expectancy at birth tends to be lower in less-developed countries compared to their more-developed counterparts (Liou *et al.*, 2020).

The term "premature mortality" holds substantial significance in assessing international and country-level performance, aiming to capture aspects of "unnecessary" or "avoidable" mortality burdens. Despite its intuitive appeal and potential policy relevance, the measurement of premature mortality poses challenges due to its latent nature. Consequently, various metrics for premature mortality are currently in use, reflecting the complexity of accurately quantifying this concept.

Calculating the burden of premature mortality involves several methods: (i) assessing the proportion of premature mortality under a chosen age threshold, (ii) determining age-standardized mortality rates within a defined age range, (iii) computing years of life lost (YLL) (Gardner & Sanborn, 1990), (iv) estimating the probability of dying within a specific age range using the life table method (Cullen, Cummins & Fuchs, 2012), and (v) comparing the standardized mortality ratio (SMR) of a study population to that of a reference population (New Mexico's Indicator-Based Information System, 2018). Among these methods, YLL and ASMR stand out as standard parameters, providing a more accurate measurement of premature mortality.

2.3.2 Years of life lost

Years of life lost or YLL is a standard parameter measurement of premature mortality. This indicator accounts for death numbers and the age at which the death occurs, giving more weight to deaths at younger ages (Romeder & Mcwhinnie, 1977;

Gardner & Sanborn, 1990). The method of calculating YLL varies from author to author. In general, two methods are commonly used to calculate YLL: i) years of potential life lost (YPLL) and ii) standard expected years of life lost (SEYLL).

YPLL was first introduced in 1941 for the tuberculosis mortality study (Dempsey, 1947). In 1971, Romeder et al. (Romeder & Mcwhinnie, 1977) refined the method of calculating YPLL as a useful mortality index for health planning, and then in 1990, the formula for YPLL was adopted by Gardner (Gardner & Sanborn, 1990). YPLL was commonly used because it was easy to calculate by subtracting the age of death from a chosen cut-off (e.g., 65, 75, or 85 years) (Gardner & Sanborn, 1990; CDC., 1993). The conventional age threshold measures of YPLL, however, do not account for deaths after the cut-off age, leading them to fail in capturing avoidable deaths at ages outside the selected age range. Furthermore, the selection of the upper age limit varies from study to study.

In 1996, the GBD study introduced SEYLL to address the issue of arbitrary age threshold selection (Murray, 1994). The SEYLL formula is based on comparing the age of death to the standard life expectancy of an individual at each age and incorporates time discounting and age weighting. Consequently, SEYLL is increasingly used as an indicator of premature mortality to calculate the mortality-associated disease burden.

2.3.3 Age standardized mortality rate

Age-standardized mortality rate or ASMR is a commonly used measure to assess premature mortality in a population. ASMR adjusts for differences in the age distribution of populations, which can vary widely between countries or regions (WHO, 2023c). By controlling for these differences, ASMR allows comparisons between populations with different age structures. ASMR can be used to monitor changes in premature mortality over time and to compare mortality rates between different

populations. It is a useful tool for identifying health disparities and evaluating the impact of public health interventions.

The age limit for calculating ASMR for premature mortality varies depending on the individual context and purpose of the analysis. WHO considers an ASMR for premature mortality between the ages of 30 and 69 years (WHO, 2018b), while some studies report an ASMR below 65 (Huisman & Bonneux, 2009) and 75 (Mackenbach *et al.*, 2015).

To calculate ASMR, age-specific mortality rates (i.e., the number of deaths within specific age groups divided by the corresponding population size) are applied to a standard population structure. The standard population structure is usually based on the age distribution of a reference population, such as a national or international standard. The age-specific mortality rates for each age group are then weighted based on the standard population structure, and the weighted rates are summed to obtain the ASMR. The ASMR is then calculated as the ratio of the expected number of deaths to the corresponding standard population size, expressed as a rate per 100,000 or 1,000 population (Naing, 2000a).

In summary, the ASMR calculation entailed two steps (Naing, 2000b): 1) calculate age-specific mortality rates. This involves calculating the number of deaths in each age group and dividing it by the corresponding population size for that age group. The result is the age-specific mortality rate. 2) Apply the standard population structure. The age-specific mortality rates are then multiplied by the corresponding proportion of the standard population in each age group. The ASMR is then calculated as the ratio of the expected number of deaths to the corresponding standard population size, expressed as a rate per 100,000 or 1,000 population. The formula for calculating ASMR is as follows:

$$ASMR = \frac{\sum (\text{age specific mortality rate} \times \text{standard population proportion})}{(\text{Standard population size})} \times 100,000$$

where, Σ represents the sum over all age groups.

2.3.4 Global burden of Premature CVD mortality

Premature mortality is a public health problem not just for CVD but for all four major categories of NCDs (CVD, cancer, chronic obstructive lung disease, and diabetes mellitus). They collectively contribute to over 80% of all premature mortality due to NCDs (WHO, 2022). Globally, 17 million premature deaths (under the age of 70) were caused by NCDs where, 38% of that was attributable to CVDs (WHO, 2021a). Furthermore, CVDs in LMICs are significant contributors to the escalating burden of NCDs, accounting for approximately 86% of all premature NCD deaths globally (WHO, 2022).

This premature mortality from NCDs results in loss of productivity and has an impact on the economy. Thus, in 2013, the WHO Global Action Plan for the Prevention and Control of NCDs 2013-2020 targeted a 25% reduction in the probability of premature death attributable to NCD by the year 2025 (WHO, 2013). Global and regional benchmarking play an important role to track the progress toward WHO Global Action Plan target. However, it remains unclear how trends in the specific type of CVD vary across population subgroups as well as the associated contributions of specific risk factors to CVD mortality. In 2015, using GBD data, Roth et al. conducted a study to project the global and regional premature cardiovascular mortality in 2025. Their study suggested if the currently observed trends in tobacco use, diabetes mellitus, obesity, and