

Forecasting subjective cognitive decline: an artificial intelligence approach using dynamic Bayesian networks

Antti Etholén, Teemu Roos, Mirja Hänninen, Ioanna Bouri, Jenni Kulmala, Ossi Rahkonen, Anne Kouvonen, Tea Lallukka

Submitted to: Journal of Medical Internet Research
on: August 02, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	38
Figures	39
Figure 1.....	40
Figure 2.....	41
Figure 3.....	42
Figure 4.....	43
Multimedia Appendixes	44
Multimedia Appendix 1.....	45
Multimedia Appendix 2.....	45
Multimedia Appendix 3.....	45
Multimedia Appendix 4.....	45

Forecasting subjective cognitive decline: an artificial intelligence approach using dynamic Bayesian networks

Antti Etholén¹ M.Sc. (Tech.), M.D.; Teemu Roos² Ph.D.; Mirja Hänninen^{1,3} M.D.; Ioanna Bouri² M.Sc; Jenni Kulmala^{4,5,6} Ph.D; Ossi Rahkonen¹ Ph.D; Anne Kouvonen^{7,8} Ph.D; Tea Lallukka¹ Ph.D.

¹ Department of Public Health University of Helsinki Helsinki FI

² Department of Computer Science University of Helsinki Helsinki FI

³ Western Uusimaa Wellbeing Services County Social and Health Care Services Espoo FI

⁴ Faculty of Social Sciences (Health Sciences) and Gerontology Research Center (GEREC) Tampere University Tampere FI

⁵ Population Health Unit Finnish Institute for Health and Welfare Helsinki FI

⁶ Division of Clinical Geriatrics, Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society Karolinska Institutet Solna SE

⁷ Faculty of Social Sciences University of Helsinki Helsinki FI

⁸ Centre for Public Health, Queen's University Belfast Royal Victoria Hospital Belfast IE

Corresponding Author:

Antti Etholén M.Sc. (Tech.), M.D.

Department of Public Health
University of Helsinki
PO BOX 20 (Tukholmankatu 8 B)
Helsinki
FI

Abstract

Background: Several potentially modifiable risk factors are associated with subjective cognitive decline (SCD). However, developmental patterns of these risk factors have not been used before to forecast later SCD. Practical tools for the prevention of cognitive decline are needed.

Objective: We examined multifactorial trajectories of risk factors, and their associations with SCD using an artificial intelligence approach to build a score that forecasts later SCD.

Methods: Five repeated surveys (2000–2022) of the Helsinki Health Study (n = 8960, 79% women, aged 40–60 at Phase 1) were used to build dynamic Bayesian networks (DBN) for estimating the odds of SCD. A score-based approach was implemented for learning DBN using the quotient normalized maximum likelihood criterion. The model was used to predict SCD based on the history of consumption of fruit and vegetables, smoking, alcohol consumption, leisure-time physical activity, body mass index, and insomnia symptoms, adjusting for sociodemographic covariates.

Results: Of the participants, 31–48% reported decline in memory, learning, and concentration in 2022. Physical activity was the strongest predictor of SCD in a 5-year interval, with an odds ratio of 0.76 (95% Bayesian credible interval 0.59–0.99) for physically active compared to inactive participants. Alcohol consumption showed a U-shaped relationship with SCD. Other risk factors had minor effects.

Conclusions: A new online risk score tool was developed that enables individuals to inspect their own risk profiles, as well as explore potential targets for interventions and their estimated contributions to later SCD. Dynamic decision heatmap was presented as a communication tool to be used at healthcare consultations.

(JMIR Preprints 02/08/2024:65028)

DOI: <https://doi.org/10.2196/preprints.65028>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>, I will be able to make my accepted manuscript PDF available to anyone.

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Forecasting subjective cognitive decline: an artificial intelligence approach using dynamic Bayesian networks

Antti Etholén¹, Teemu Roos², Mirja Hänninen^{1,3}, Ioanna Bouri², Jenni Kulmala^{4,5,6}, Ossi Rahkonen¹,
Anne Kouvonen^{7,8}, Tea Lallukka¹

¹ Department of Public Health, P.O. Box 20 (Tukholmankatu 8 B), 00014 University of Helsinki, Finland

² Department of Computer Science, P.O. Box 68, University of Helsinki

³ Western Uusimaa Wellbeing Services County, Social and Health Care Services, P.O. Box 33, 02033 Espoo, Finland

⁴ Faculty of Social Sciences (Health Sciences) and Gerontology Research Center (GEREC), Tampere University, Arvo Ylpön katu 34, 33520 Tampere, Finland

⁵ Population Health Unit, Finnish Institute for Health and Welfare, POB 30, 00271 Helsinki

⁶ Division of Clinical Geriatrics, Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Solnavägen 1, 171 77, Solna, Sweden

⁷ Faculty of Social Sciences, University of Helsinki, P.O. Box 54, 00014 University of Helsinki, Finland

⁸ Centre for Public Health, Queen's University Belfast, Royal Victoria Hospital, Belfast, BT12 6BA, Northern Ireland

***Corresponding author:**

Antti Etholén, M.D., M.Sc. (Tech.)

Department of Public Health, P.O. Box 20

FI-00014 University of Helsinki, Finland

email: antti.etholen@helsinki.fi, tel. +358 44 510 5010

Contact details

Antti Etholén, M.D., M.Sc. (Tech.), antti.etholen@helsinki.fi, 0000-0002-0238-6838

Teemu Roos, Ph.D., teemu.roos@helsinki.fi, 0000-0001-9470-3759

Mirja Hänninen, M.D., mirja.hanninen@helsinki.fi, 0000-0002-6807-7333

Ioanna Bouri, M.Sc., ioanna.bouri@helsinki.fi, 0009-0002-6807-2951

Jenni Kulmala, Ph.D., jenni.kulmala@tuni.fi, 0000-0003-0194-5551

Ossi Rahkonen, Ph.D., ossi.rahkonen@helsinki.fi, 0000-0002-7202-3274

Anne Kouvonen, Ph.D., anne.kouvonen@helsinki.fi, 0000-0001-6997-8312

Tea Lallukka, Ph.D., tea.lallukka@helsinki.fi, 0000-0003-3841-3129

Abstract

Background

Several potentially modifiable risk factors are associated with subjective cognitive decline (SCD). However, developmental patterns of these risk factors have not been used before to forecast later SCD. Practical tools for the prevention of cognitive decline are needed.

Objective

We examined multifactorial trajectories of risk factors, and their associations with SCD using an artificial intelligence approach to build a score that forecasts later SCD. More specifically, our aim was to develop a new risk score to facilitate personalized risk assessment and intervention planning. In addition, we aimed to validate SCD against register-based dementia diagnoses and dementia-related medications.

Methods

Five repeated surveys (2000–2022) of the Helsinki Health Study ($n = 8960$, 79% women, aged 40–60 at Phase 1) were used to build dynamic Bayesian networks (DBN) for estimating the odds of SCD. The model structure was developed using expert knowledge and automated techniques, implementing a score-based approach for learning DBN with the quotient normalized maximum likelihood criterion. DBNs allow for modeling complex, non-linear relationships among variables and handling missing data efficiently. The developed model was used to predict SCD (memory, learning, and concentration) based on the history of consumption of fruit and vegetables, smoking, alcohol consumption, leisure-time physical activity, body mass index, and insomnia symptoms, adjusting for sociodemographic covariates. Model performance was assessed using 5-fold cross-validation to calculate the area under the receiver operating characteristic curve (AUC). Bayesian credible intervals were used to quantify uncertainty in model estimates.

Results

Of the participants, 31–48% reported decline in memory, learning, and concentration in 2022. Physical activity was the strongest predictor of SCD in a 5-year interval, with an odds ratio of 0.76 (95% Bayesian credible interval 0.59–0.99) for physically active compared to inactive participants. Alcohol consumption showed a U-shaped relationship with SCD. Other risk factors had minor effects. Moreover, our validation confirmed that SCD has prognostic value for diagnosed dementia, with individuals reporting memory decline being over three times more likely to have dementia in 2017 (aged 57–77 years), and this risk increasing to more than five times by 2022 (aged 62–82 years). The receiver operating characteristic (ROC) curve analyses further supported the predictive validity of our outcome, with an area under the curve (AUC) of 0.78 in 2017 and 0.75 in 2022.

Conclusion

A new risk score tool was developed that enables individuals to inspect their own risk profiles, as well as explore potential targets for interventions and their estimated contributions to later SCD. By integrating AI-driven predictive modeling, the new tool can aid healthcare professionals in providing personalized prevention strategies. Dynamic decision heatmap was presented as a communication tool to be used at healthcare consultations. Our findings suggest that early identification of individuals with SCD could improve targeted intervention strategies to be used to reduce dementia risk. Future research should explore the integration of AI-based risk prediction models into clinical workflows and assess their effectiveness in guiding lifestyle interventions to mitigate SCD and dementia.

Keywords: artificial intelligence; dementia; aging; smoking; alcohol consumption; leisure-time physical activity; consumption of fruit and vegetables; body mass index; insomnia symptoms



Introduction

Healthy aging is crucial for maintaining individual well-being and managing the capacity of health and social care systems [1]. The economic impact of aging is profound, with health and social care costs surging globally [2,3]. In addition, population aging significantly increases the global burden of disease [4]. This situation highlights the urgent need for cost-effective preventive healthcare strategies. Despite its critical role, preventive healthcare is often underutilized, constrained by limited medical consultation times and the prioritization of symptomatic patients. [5] Therefore, there is a pressing need for innovative, efficient, and cost-effective strategies to enhance preventive measures. Early identification of risky lifestyles already in midlife is essential; it can delay the onset of health deterioration in later years and significantly improve the quality of life by adding healthy years to life. [6]

Cognitive functioning is a crucial factor for functional ability and well-being in later life [7]. Subjective cognitive decline (SCD) — self-perceived worsening of cognitive abilities, such as memory, learning or concentration, despite normal performance on objective cognitive tests — is increasingly recognized as one of the potential early warning signs of dementia [8,9]. Risk factors for cognitive decline include high age, ApoE4 genotype, cardiovascular and metabolic conditions, mental health issues, low education, and unhealthy lifestyle factors [7,10]. Since the underlying cause of cognitive decline is still not fully understood, there are currently no curative treatments, and thus early detection and management of their risk factors are critical [11]. While traditional cognitive assessments remain valuable, innovative approaches are emerging to both assess and intervene in cognitive decline. Recent work has demonstrated the efficacy of technology-based interventions such as virtual reality for enhancing cognitive functioning in those already experiencing mild cognitive impairment [12].

Traditional predictive models, such as logistic regression and Cox proportional hazards models, often assume linear relationships and may struggle to capture complex, time-dependent risk interactions. Microsimulation models simulate individual trajectories but rely on predefined transition structures, which may limit flexibility. In contrast, dynamic Bayesian networks (DBNs) offer a data-driven approach that models probabilistic dependencies over time, providing greater adaptability in uncovering latent patterns and health trajectories. Building on these advantages, this study applies advanced artificial intelligence (AI) techniques, particularly dynamic Bayesian networks (DBN), to robustly analyze complex data, reveal latent patterns over time, and identify risk-related patterns for SCD. This approach promises a more personalized and proactive identification of health trajectories. [13–15]

The main aim of this study is to produce new information about multifactorial trajectories of risk factors and their associations with SCD using an artificial intelligence approach, developing a new score to forecast SCD. Furthermore, the aim is to validate an SCD instrument with objective register data on dementia. The main aim is divided into four objectives:

1. Developing and applying AI models: This objective includes using advanced AI technologies to develop and assess the DBN for risk estimation and modeling risk factor trajectories. A key element is the creation of the Helsinki Health Study (HHS) score for predicting SCD using indicators such as smoking, alcohol consumption, leisure-time physical activity (LTPA), consumption of fruit and vegetables, body mass index (BMI), and insomnia symptoms, adjusting for sociodemographic factors.
2. Using the HHS score to predict SCD: This objective focuses on providing examples of how the HHS score can help individuals assess their health risks and make informed decisions.
3. Creating a new tool: The final objective is to build an interactive online calculator that allows users to enter their risk factors and estimate changes in SCD over five years based on chosen

intervention.

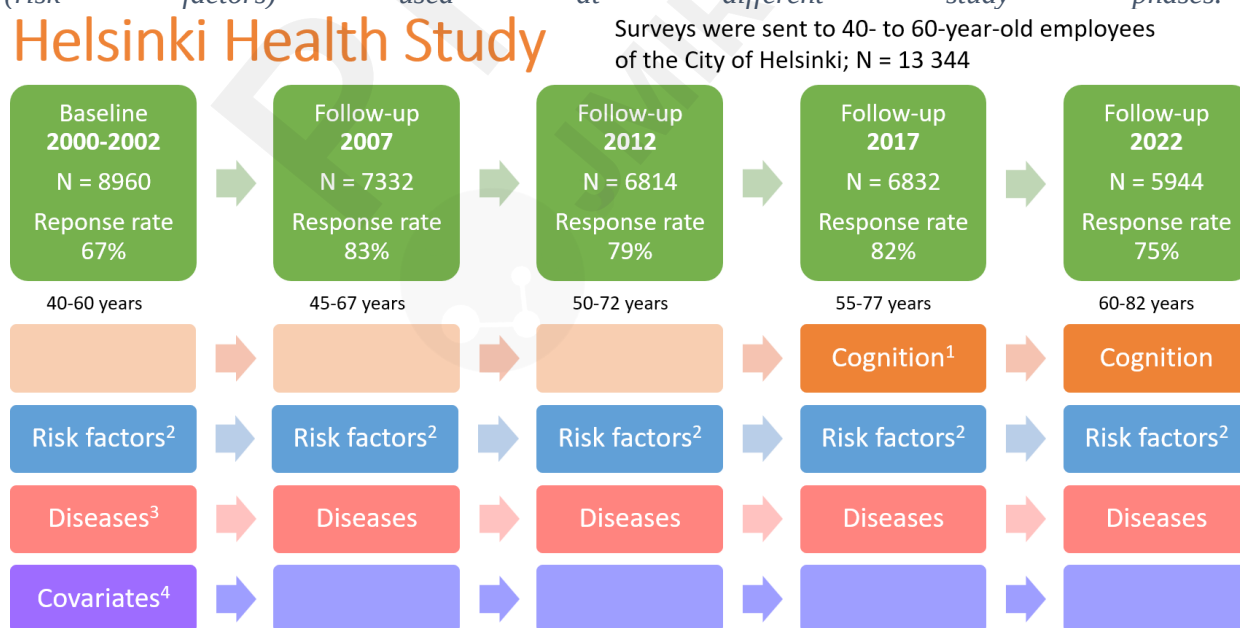
4. Validating SCD against objective register-based data on dementia diagnoses and dementia medications, for the first time. This objective helps confirm the results and provides new information about the reliability and validity of survey-based outcome for future use as well.

Methods

Data

Data for this study were obtained from the Helsinki Health Study (HHS) (Figure 1). [16] The initial survey was conducted between 2000 and 2002, targeting 40- to 60-year-old employees of the City of Helsinki, Finland (Phase 1, n=8,960, response rate: 67%). Follow-up surveys were administered to all Phase 1 participants in 2007 (Phase 2, response rate: 83%), 2012 (Phase 3, response rate: 79%), 2017 (Phase 4, response rate: 82%), and 2022 (Phase 5, response rate: 75%). The final analytic sample comprised 8960 individuals.

Figure 1. Phases of the Helsinki Health Study cohort: outcome (subjective cognitive decline) and exposures (risk factors) used at different study phases.



¹ Subjective cognitive decline (memory, learning and concentration)

² Risk factors: consumption of fruit and vegetables, smoking, alcohol consumption, physical activity, body mass index, and insomnia symptoms

³ Diseases: hypertension, high cholesterol, diabetes, pain status and psychiatric conditions (depression, anxiety, and other mental disorders)

⁴ Sociodemographic covariates: age, gender, education

Measurement

Risk

factors

Risk factors[7,17] were categorized following the current national guidelines[18] and were used as time dependent variables: consumption of fruit and vegetables (daily consumption of both, daily consumption of either, non-daily consumption), smoking (never smoker, ex-smoker, and smoker), alcohol consumption (no consumption; moderate consumption among men 1–13 units of alcohol at least once a week, women: 1–6 units; high consumption among men at least 14 units, women 7 units, very high-risk consumption among men at least 23 units, women at least 12 units) and LTPA (physically inactive: metabolic equivalent (MET) < 14 vs physically active: MET ≥ 14), BMI (underweight less than 18.5 or recommend healthy weight 18.5–24.9, overweight 25.0–29.9, obesity 30+)[19] and insomnia symptoms (<4 nights/month, 4–14 nights/month, >14 nights/month).

Subjective cognitive decline

Three variables were used to measure SCD. The participants were asked to report ‘How well my memory works’, ‘How well embracing and learning new things goes for me’, ‘Normally I can concentrate on something’ [17,20]. The instrument is included in the TOIMIA database, which is maintained by the Finnish Institute for Health and Welfare [21]. The items were used to define three different aspects of self-rated cognitive functioning: memory, learning, and concentration. Participants rated themselves on a five-point scale (‘very poorly’, ‘poorly’, ‘satisfactorily’, ‘well’, ‘very well’). The scales were re-categorized as good (including ‘well’ and ‘very well’) and declined (including ‘satisfactorily’, ‘very poorly’ and ‘poorly’). First register-based validation of the measure is available in **Appendix**.

Sociodemographic covariates

Age was utilized as a dynamic variable (time-dependent) and segmented into the closest age groups (40–80 years with 5-year intervals). Additionally, other variables served as static variables in the DBN model: gender (women, men; Phase 1) and education (basic, secondary, higher; Phase 1).

Health factors and physician diagnosed diseases

Several health factors were included to underscore their potential associations with SCD. These factors encompassed current pain status (no, yes), and the following physician-diagnosed conditions: hypertension, high serum cholesterol, diabetes, and mental disorders (depression, anxiety disorder, other mental disorder).

Statistical

analyses

AI approach in the development of the HHS score

We developed the HHS score using DBN, allowing detailed analysis of relationships among covariates, variables, and outcomes over time segments [22,23]. DBNs, structured as directed acyclic graphs, include nodes representing variables and edges denoting probabilistic dependencies. [24] This structure offers a transparent analytical framework, enhancing model predictability and accuracy. The time-homogeneous DBN maintains consistency while allowing influence from previous time steps. We modeled temporal progression by age, with each step representing five years, accurately depicting health-related variable progression.

The statistical analyses and model building were performed using the C# programming language [25]

with the R.NET library [26] functioning as an in-process bridge to integrate R version 4.2.1 [27] with the bnlearn-package [28]. Probabilistic inferences within the network were facilitated by the ML.NET probabilistic package [29].

Data preparation

Outliers were treated as missing values. We generated descriptive contingency tables to illustrate missing data and the sizes of each variable group across Phases 1–5 (see **Table 1**). Due to the infrequency of missing covariate data and the capability of Bayesian networks to handle such values, we did not perform any imputations [30]. Moreover, the full dataset could be used in the model, as each node's conditional probability table (CPT) was estimated based on the available cases for that specific variable, without requiring complete data across all variables. All categorical variables were treated as discrete random variables, with domains that are mutually exclusive and exhaustive.

Table 1

Background characteristics and Dynamic Bayes Network node names and parent nodes. The study population is based on survey data from the Helsinki Health Study (2000–2022). Baseline (Phase 1) study population consisted of 40- to 60-year-old employees of the City of Helsinki, Finland.

	Phase 1 2000- 2002 (N=896 0)	Phase 2 2007 (N=73 32)	Phase 3 2012 (N=68 08)	Phase 4 2017 (N=68 31)	Phase 5 2022 (N=59 50)
--	---	---	---	---	---

SOCIODEMOGRAPHIC FACTORS

Gender

Men	20.0%
Women	80.0%

Age group

40 years	20.1%				
45 years	21.2%	19.1%			
50 years	21.7%	20.4%	18.3%		
55 years	24.9%	21.7%	20.5%	19.3%	
60 years	12.1%	26.0%	22.0%	21.4%	20.4%
65 years		12.7%	26.6%	22.2%	22.7%

70 years	12.7%	25.5%	22.2%
75 years		11.6%	24.4%
80 years			10.3%

Education

Basic	42.7%
Secondary	31.4%
Higher	25.9%
No answer	79

RISK FACTORS**Smoking**

Never	50.9%	51.8%	50.4%	47.0%	47.1%
Ex-smoker	24.3%	29.3%	34.1%	39.0%	42.8%
Current	24.8%	18.9%	15.5%	14.0%	10.1%
No answer	442	393	692	993	934

Alcohol consumption

No consumption	7.0%	8.5%	10.5%	14.2%	18.3%
Moderate consumption	76.5%	72.6%	73.5%	71.9%	71.6%
High consumption	9.7%	10.2%	8.9%	8.0%	5.7%
High risk consumption	6.7%	8.8%	7.1%	5.9%	4.4%
No answer	387	265	260	206	279

Leisure-time physical activity

Physically inactive (MET < 14)	25.2%	22.7%	26.5%	24.2%	24.5%
Physically active (MET > 14)	74.8%	77.3%	73.5%	75.8%	75.5%
No answer	91	76	97	112	54

Fruit and vegetable consumption

Non-daily consumer	23.9%	20.0%	18.8%	21.6%	15.4%
Daily consumer of either	29.3%	28.3%	27.3%	25.0%	24.0%
Daily consumer both	46.8%	51.7%	53.9%	53.4%	60.6%
No answer	87	129	133	173	139

Body mass index

Underweight -18.5 or recommend healthy weight 18.5+	51.4%	46.1%	42.4%	40.4%	40.4%
Overweight 25+	34.2%	35.7%	36.7%	37.8%	37.2%
Obesity +30	14.4%	18.2%	20.9%	21.9%	22.4%
No answer	107	80	102	88	87

Insomnia symptoms

<4 nights/month	49.1%	42.7%	44.2%	42.7%	42.4%
4-14nights/month	31.6%	32.2%	30.9%	29.9%	30.4%
>14 nights/month	19.3%	25.0%	24.9%	27.4%	27.2%

No answer	744	613	619	976	690
-----------	-----	-----	-----	-----	-----

SUBJECTIVE COGNITIVE DECLINE

Memory

Good	67.6%	68.6%
------	-------	-------

Declined	32.4%	31.4%
----------	-------	-------

No answer	69	85
-----------	----	----

Learning

Good	52.3%	52.1%
------	-------	-------

Declined	47.7%	47.9%
----------	-------	-------

No answer	117	71
-----------	-----	----

Concentration

Good	69.5%	69.0%
------	-------	-------

Declined	30.5%	31.0%
----------	-------	-------

No answer	113	62
-----------	-----	----

HEALTH FACTORS AND PHYSICIAN DIAGNOSED DISEASES

Pain

No	55.4%	55.3%	57.7%	56.3%	55.7%
----	-------	-------	-------	-------	-------

Yes	44.6%	44.7%	42.3%	43.7%	44.3%
-----	-------	-------	-------	-------	-------

No answer	205	226	160	337	239
-----------	-----	-----	-----	-----	-----

Hypertension

No	77.1%	67.1%	61.4%	48.5%	40.2%
----	-------	-------	-------	-------	-------

Yes	22.9%	32.9%	38.6%	51.5%	59.8%
-----	-------	-------	-------	-------	-------

No answer	95	125	109	1400	933
-----------	----	-----	-----	------	-----

High cholesterol

No	78.8%	65.2%	56.8%	54.0%	43.8%
----	-------	-------	-------	-------	-------

Yes	21.2%	34.8%	43.2%	46.0%	56.2%
-----	-------	-------	-------	-------	-------

No answer	86	80	35	1764	1310
-----------	----	----	----	------	------

Diabetes

No	97.0%	93.2%	88.6%	82.8%	79.4%
----	-------	-------	-------	-------	-------

Yes	3.0%	6.8%	11.4%	17.2%	20.6%
-----	------	------	-------	-------	-------

No answer	1071	1652	1464	2173	1889
-----------	------	------	------	------	------

Mental disorders

No	81.5%	77.1%	78.9%	79.4%	79.6%
----	-------	-------	-------	-------	-------

Yes	18.5%	22.9%	21.1%	20.6%	20.4%
-----	-------	-------	-------	-------	-------

No answer	858	1435	1365	2193	1894
-----------	-----	------	------	------	------

Implementing the DBN model

Initially, we established the model structure using expert knowledge and automated techniques. Temporal variables (age group, consumption of fruit and vegetables, smoking, alcohol consumption, LTPA, BMI and insomnia symptoms) were hardwired between time slices to represent development paths. We then used a score-based quotient normalized maximum likelihood (qNML) criterion algorithm to identify additional relationships within the same time slice (horizontal effects) between nodes. qNML is an information-theoretic model selection criterion similar to the commonly used BIC criterion [31]. It identifies the most informative parent sets for each variable while penalizing model complexity, qNML has been found to favor models that, while parsimonious, tend to achieve higher predictive performance than models selected by BIC for moderate and large sample sizes [32]. In the qNML algorithm, the nodes were ordered as follows: age group, gender, education, consumption of fruit and vegetables, smoking, alcohol consumption, LTPA, insomnia symptoms, BMI, hypertension, high cholesterol, diabetes, psychiatric disorders, current pain status, concentration, memory, and learning. The decomposable qNML algorithm facilitates separate structural learning for each node and its potential parents. To manage complexity, we limited each node to a maximum of three parents. Variables were organized so causative variables preceded effects, but constructing a causal network was not our intent. [33] We assessed combinations of up to three parents for each node, selecting the configuration with the highest qNML score. Next, we derived conditional probability distribution tables (CPTs) for each node using our complete dataset. To address the zero-cell problem in probability estimation, we applied the Bayesian smoothing algorithm with a Dirichlet prior ($1/2, \dots, 1/2$) to each CPT.

Inference using the model

The use of BN offers significant advantages, primarily the ability to perform inference based on specified evidence and events. Age is the initial parameter in this dynamic model. For the initial time

slice, the CPTs represent prior probabilities from our dataset. This inference mechanism supports a scoring calculator where users input specific data into the network. This allows calculation of event probabilities, such as the likelihood of memory decline or probabilities related to the longitudinal development of variables. Here, 'evidence' refers to user-provided data used to infer the HHS score based on a pre-trained DBN.

The Expectation Propagation algorithm, a deterministic method, was used for inference, yielding approximate results in discrete cases [34]. In inference, information contained in the evidence provided by the user is propagated through the network one variable at a time, updating the probability distribution of each variable by conditioning on the evidence. We calculate the full posterior distribution for each variable within each time slice (t) to maintain uncertainty and probabilistic relationships among variables. Results are presented with Bayesian credible intervals, which inherently account for uncertainty and ensure robust estimation of relationships while preserving probabilistic dependencies over time. When transitioning from one time slice to the next, we use the Maximum A Posteriori (MAP) estimates of the variables from the previous time slice ($t-1$) as inputs for calculating the current time slice (t) to reduce computational complexity and prevent the over-propagation of uncertainty. Cross-correlations and autocorrelations among output variables were also calculated for randomly sampled 100 individuals from our dataset, based on demographic and risk factors.

Evaluation metrics

We validated our model using 5-fold cross-validation and evaluated classification performance by calculating the area under the receiver operating characteristic curve (AUC) [35].

Decision heatmap

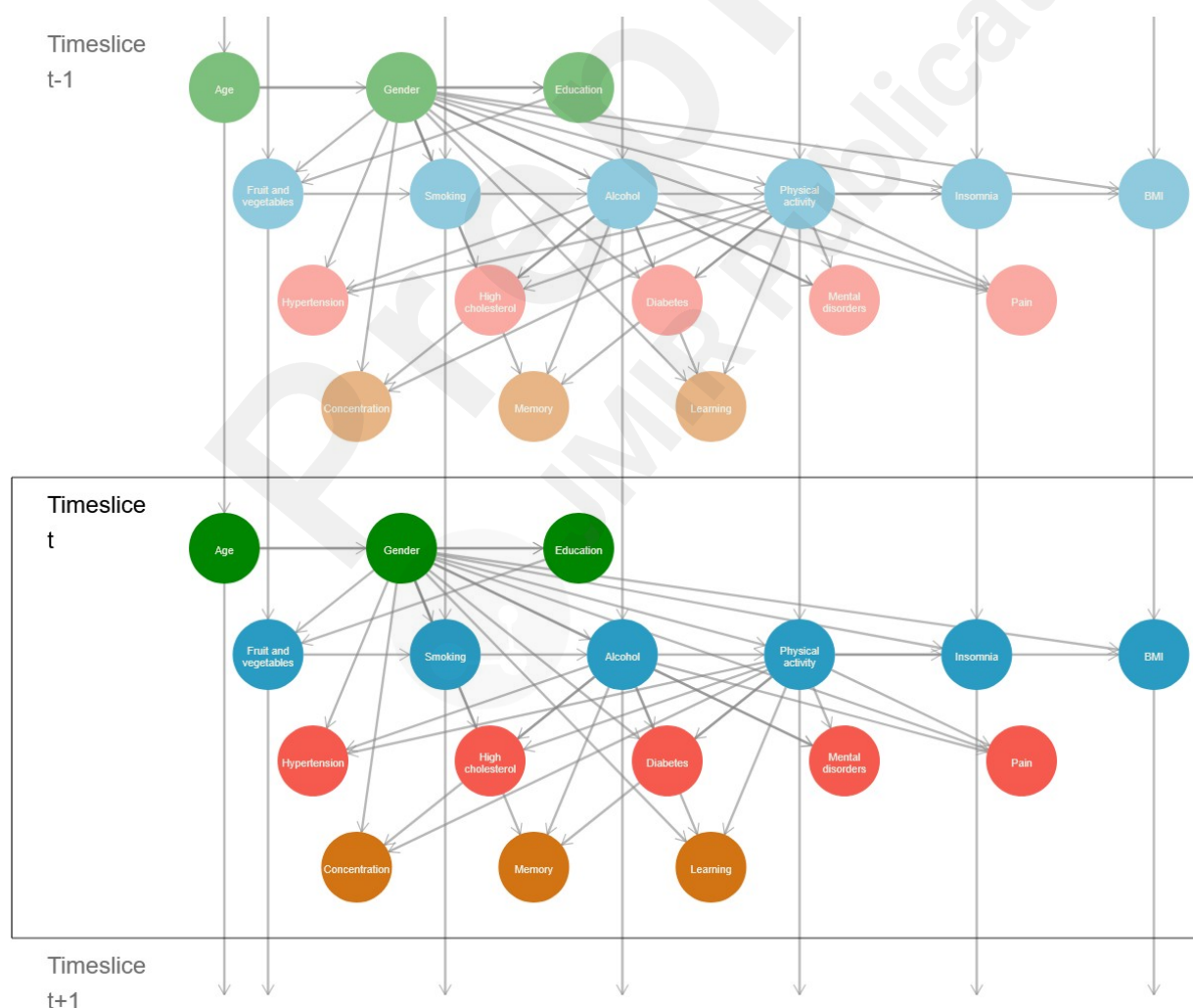
We introduce a dynamic decision heatmap as a communication tool for instance in patient consultations. This heatmap displays the current state of a patient's lifestyle and presents a decision matrix for each modifiable risk factor. It illustrates the effects of potential changes, projected five years into the future, showing risk changes derived from the DBN. When selecting a target intervention, the heatmap initiates pre- calculation in the DBN using the chosen intervention as evidence, generating new estimate for SCD to facilitate informed decision-making.

Results

Of the participants, 31% reported decline in memory, 48% in learning abilities, and 31% in concentration (**Table 1**). Non-daily consumption of fruit and vegetables was reported by 15–24% of the participants, while 48–53% were current or former smokers. Additionally, 10–19% of the participants engaged in high or high-risk alcohol consumption. Physical inactivity was reported by 23–27% of the participants, and 54–60% were with overweight or obesity. Finally, insomnia symptoms occurring 14 nights per month were reported by 19–27% of the participants.

The structure of the DBN model, learned using qNML, is illustrated in **Figure 2**. The most informative parent nodes for the nodes representing memory, learning, and concentration were gender, alcohol consumption, and LTPA. These same parent nodes were selected for the nodes representing pain, hypertension, high cholesterol, diabetes, and mental disorders diagnosed by a physician.

Figure 2. The HHS score dynamic Bayesian network model for predicting subjective cognitive decline (SCD) based on the expert knowledge and score-based quotient normalized maximum likelihood criterion algorithm. The flow (possible parents of current node) of network is from left-to-right and from up-to-bottom. Definitions for nodes: Age=Age group (5-year intervals), Gender (woman, man), Education (basic, secondary, higher), Fruit and vegetables = consumption of fruit and vegetables (daily consumption of both, daily consumption of either, non-daily consumption), Smoking (never smoker, ex-smoker, and smoker), Physical activity =leisure-time physical activity (physically inactive: metabolic equivalent (MET) < 14 vs physically active: MET ≥ 14), BMI=body mass index (underweight less than 18.5 or recommend healthy weight 18.5–24.9, overweight 25.0–29.9, obesity 30+), Insomnia=insomnia symptoms (<4 nights/month, 4–14 nights/month, >14 nights/month), Hypertension (no, yes), Highcholesterol=high cholesterol(no, yes), Diabetes (no, yes), Mental disorders (no, yes; including depression, anxiety disorder, other mental disorder), Pain=current pain status (no, yes), Concentration (good, declined), Memory (good, declined), Learning (good, declined). Total training size N = 8960 based on five repeated surveys (2000–2022) of the Helsinki Health Study.



Analysis of the CPTs presented in **Table S1** indicates that LTPA consistently reduces the probability

of SCD. Specifically, the average relative risk reduction of experiencing declined memory when transitioning from inactive to active physical states is 26.9% for men and 24.8% for women, across different levels of alcohol consumption. Additionally, the relationship between alcohol consumption and SCD follows a U-shaped curve, with both non-consumption and high-risk consumption exerting greater effects than moderate or elevated consumption. Manual testing of the network, corroborated by data from our sampled participants, supports these findings.

The user interface of the HHS score calculator is presented in **Figure 3**. The inputs for the calculator are set as evidence for Bayesian inference. The probabilities for each output, including developmental paths and risks, can be read from the inferred model.

Figure 3. The HHS score user interface allows users to their input current risk factors, demographic variables, and potential targets for five years later, generating trajectories of risk factors and risks associated with subjective cognitive decline (SCD) and health factors. For example, in this case the current evidence = 55 years, woman, physically inactive and target evidence = 55 years, woman, physically inactive (55 years), physically active (60 years). The risk of declined memory when 60-year-old based on current evidence = 40.1% and based on target evidence = 33.9% having odds ratio OR 0.76 (95% Bayesian credible interval 0.59–0.99). MET= metabolic equivalent.



We selected an individual example from our sample to demonstrate the use of the HHS score calculator (**Figure 3**). This individual is a 55-year-old woman with secondary education, who is not eating fruit and vegetables daily, is a current smoker, does not consume alcohol, is physically inactive, is with overweight and has high insomnia symptoms. In a consultation with a health care

professional, the decision heatmap (**Figure 4**) can be utilized to communicate with the individual and propose personalized lifestyle modifications. For instance, a strategy focused on increasing LTPA over the next five years appears to be particularly effective. The outcomes indicate that, for example, the likelihood of subjective memory decline decreases from 40.2% to 33.9%, an absolute likelihood reduction of 6.3 percentage points, and an odds ratio of 0.76 (95% Bayesian credible interval 0.59–0.99) (**Figure 3**). Considering that continuing the current lifestyle is estimated to lead to a 49.2% likelihood of memory decline, an intervention resulting in a reduction of 15.3 percentage points is estimated to decrease this likelihood to 33.9% (**Figure 4**).

Figure 4. The first step in the decision heatmap to decrease self-reported memory decline of a 55-year-old woman with a secondary education. The decision matrix shows effects of possible changes 5 years from now based on the current state. When the first choice is made, the evidence is changed and the new values to matrix are updated making this decision heatmap dynamic.

Current state (now)		Effects of possible changes (or their lack) from current to later self-rated memory decline five years later					
Smoking	Current	Smoking	Alcohol consumption	Physical activity	Fruit and vegetables	Body mass index	Insomnia
Alcohol consumption	No	Never (no applicable)	No 0%	Inactive 0%	None 0%	Normal -0.5%	None +0.3%
Physical activity	Inactive	Ex-smoker -0.3%	Moderate -10.8%		Either -1.3%	Overweight 0%	Medium +0.1%
Fruit and vegetables	None		High -18.3%	Active -15.3%	Both -1.8%		
Body mass index	Overweight	Current 0%	High risk -5.6			Obesity -0.5%	High 0%
Insomnia	High						

We further found statistically significant cross-correlations between decline in SCD and various commonly reported health conditions, including hypertension ($p < 0.001$), high serum cholesterol ($p < 0.001$), and diabetes ($p < 0.001$), based on a sample at the initial time step (t) (**Table S2**). Mental disorders and pain showed minor contribution to the decline in SCD. Following intervention, the effects intensified at time step ($t+1$) and subsequent autocorrelation analysis indicated stabilization from ($t+1$) to ($t+3$).

The model's predictive power measured by the AUC metric was moderate (cross-validated AUC scores were in the range 0.52–0.65) (**Table S3**). This indicates that the model provides meaningful

estimates of average risk differences across lifestyle profiles.

Discussion

Main findings

We demonstrated how an AI approach with DBN can be utilized to build a risk score for forecasting SCD, effectively handling complex interactions between different factors, and providing clear, interpretable outcomes. We found that LTPA appears to have the most consistent associations with subsequent SCD. In addition, we discovered that SCD and disease risks such as diabetes, hypertension and high serum cholesterol were highly correlated.

Interpretation

The nature of SCD is multifactorial and may arise from normal aging as older adults are more likely to monitor their memory and raise concerns, be reversible, or progressively lead to dementia [8]. SCD is prevalent in aging populations, and prevention of dementia is a global priority for the World Health Organization (WHO) [36,37]. Additionally, a lifestyle intervention program has been found to have beneficial effects for preventing cognitive decline in at-risk elderly people [38]. This highlights the importance of the preventive HHS score, which addresses multifactorial issues through advanced AI technique DBN to analyze complex data and identify emerging risk groups already in midlife.

Our study shows the complex associations between alcohol use, smoking, LTPA, consumption of fruit and vegetables, body mass index, and the history of insomnia symptoms and SCD. We found that changing these potentially modifiable factors could decrease the overall likelihood of SCD. Notably,

LTPA emerged as the most significant predictor in our model, decreasing the risk of SCD. Numerous interventions have aimed to enhance cognitive functioning, and for example, a systematic review by Barha et al. has shown that especially aerobic exercise has a positive impact on cognition (standardized mean difference 0.85 SD higher compared to controls) [39]. This effect size aligns with previous cohort studies, where high physical activity in mid-to-late life has been associated with up to a 38% lower risk of SCD [40], and physical inactivity has been linked to increased odds of SCD [41]. This finding aligns with the well-documented benefits of LTPA on cognitive function [42,43]. Other studies also corroborate that physical inactivity is a significant risk factor for SCD [44,45]. Many studies demonstrate the preventive importance of physical activity in achieving healthy aging and preventing frailty [46]. Increased physical activity can improve cerebral blood flow, enhance synaptic plasticity, facilitate cellular cleanup processes, decrease beta-amyloid levels, improve mental health, and reduce stress [47]. The novelty of this study lies in a new tool that assesses how changes in risk factors impact SCD during follow-up and how increased physical activity, for example, can reduce its likelihood.

The observed correlation between SCD and other disease risks may indicate shared underlying mechanisms or common risk factors. This suggests that interventions targeting cardiovascular risk factors could concurrently improve SCD [48]. Cognitive decline is also influenced by genetic factors, such as the ApoE4 genotype, as well as health conditions, which may interact with modifiable lifestyle factors [7]. Future studies should explore these interactions to better understand their role in the SCD risk.

Practical applications

Given the shared risk factors, insights into SCD can provide valuable indications about potential

progression to dementia. However, since our findings are based on observational data, the projected associations illustrated by the decision heatmap and the HHS score calculator should not be interpreted as direct evidence of intervention efficacy without confirmatory intervention studies. While the longitudinal design, in which risk factors were assessed prior to the onset of SCD, supports the possibility of a causal relationship, it does not confirm that modifying these risk factors would influence the outcome. For that, randomized controlled trials are needed. However, follow-up time in RCT is often short and the long-term effects of interventions are difficult to capture, particularly for health outcomes that develop slowly (such as dementia). Furthermore, it is possible to produce findings highly like RCTs even with observational data [49]. Prior to broader implementation, it is crucial to conduct an intervention study focusing on the identified risk factors to determine if risk factor modifications can indeed lower the risk of SCD or dementia in line with the results suggested by the algorithm, and what is the adherence to different lifestyle interventions in the long run.

The HHS score partially overlaps with existing dementia risk scores in terms of included lifestyle-related predictors. For example, the CAIDE score incorporates BMI and physical inactivity, but not smoking, alcohol, diet, or sleep-related factors [50], while the LIBRA score covers most lifestyle components—smoking, alcohol, diet, and physical activity—but does not include insomnia [51]. Unlike these additive risk models estimating long-term dementia risk, the HHS score uses a dynamic Bayesian network to predict SCD based on longitudinal data and interdependencies between variables. This approach allows for more individualized, time-sensitive predictions. A recent review emphasized the need for more dynamic, personalized tools for cognitive risk estimation, particularly those accounting for time-varying exposures [51].

In clinical settings, our score calculator could assist in assessing the probabilities of SCD. Additionally, our decision heatmap supports health care professionals when discussing how lifestyle

changes could impact health outcomes with their patients. Our methodology integrates decision theory—combining elements of probability theory and utility theory—to evaluate the effectiveness of preventive measures [52]. For example, our DBN identified a U-shaped curve in alcohol consumption, similar to as a large meta-analysis [53]. This finding exemplifies how latent effects can influence risk assessments and should be interpreted with respect to probabilities. This integrated approach ensures that recommendations are personalized and rooted in a comprehensive assessment of each patient's unique health profile and needs. To support real-world clinical use, potential implementation barriers—such as integration with electronic health records, time constraints during appointments, and provider training—should also be considered. However, given the moderate predictive power (AUC 0.52–0.65), the HHS score should currently be considered a supportive tool rather than a definitive diagnostic instrument. Further external validation and refinement are needed before clinical implementation.

Limitations and strengths of the study

Our study had some limitations. First, our attrition analysis [17] indicated that individuals with SCD were more likely to drop out in the subsequent phase. Second, our results cannot be directly generalized to the general population, as the cohort was female-dominated and composed of individuals who were employees of the City of Helsinki at inclusion. While this homogeneity reduces confounding, it may limit external validity. The HHS score may perform differently in populations with more diverse socioeconomic backgrounds, in general population, different occupations, depending on healthcare access, or in male-dominated populations. Importantly, gender was identified as an informative parent node for several other variables in the model. This means that, due to the underrepresentation of men in the cohort, the conditional probability tables (CPTs) for male-specific paths in the network were estimated from fewer observed values, potentially

affecting the model's performance and stability in male-dominated or gender-balanced populations. Future studies should validate the model in more heterogeneous samples.

Despite these limitations, our study had also significant strengths. Firstly, our over 20-year longitudinal cohort used identical questionnaires across all five phases, achieving moderate to excellent response rates for a broad range of sociodemographic and health-related variables. Non-response analyses confirmed the data accurately represented the target population.[16] Secondly, we used an AI approach to better account for the complex dynamics and interactions among our variables. DNB is an ideal model for our calculator because of its ability to process various types of data. It can effectively manage cases where some nodes are in an undefined state, labeled as 'NA', without compromising its functionality. This model simplifies the complexity typically encountered in networks by avoiding connections between every pair of variables. Instead, it links variables only across consecutive time slices. Unlike deep learning, which often requires large datasets and can suffer from reduced interpretability, DBNs are well-suited for modeling structured relationships in smaller datasets. Additionally, DBNs allow for the simultaneous analysis of dynamic interactions among multiple variables over time, rather than assuming linear relationships as in traditional regression models. Additionally, the DBN model is characterized by its transparency, allowing users to inspect and understand the internal processes. Moreover, it can handle incomplete data sets without requiring data imputation to fill missing values. However, DBNs are not without limitations. They can be sensitive to data quality, particularly when dealing with small sample sizes or noisy data, and they may be prone to overfitting if not properly regularized. Thirdly, a key strength of the study is that we were able to conduct a first validation of the SCD against objective register data on dementia (medication data and hospitalizations). Our validation demonstrated a strong association between SCD and clinically verified dementia diagnoses, with individuals reporting memory decline being over three times more likely to have dementia in 2017 (aged 57–77 years), and this risk

increasing to more than five times by 2022 (aged 62–82 years). Additionally, declines in learning and concentration showed similar trends, reinforcing the role of SCD as an early indicator of later dementia risk. The receiver operating characteristic curve analyses further confirmed its predictive validity, with an area under the curve of 0.78 in 2017 and 0.75 in 2022. Our measurement of SCD further has good internal consistency and consistency remained between Phase 4 (2017) and Phase 5 (2022). The consistent trends observed across Phases suggest that the measure reliably captures the associations between SCD and many health-related factors and dimensions of health functioning. These findings confirm that SCD is a valid and reliable outcome measure, with significant prognostic value for dementia (**Appendix**). This helped confirm our results and supports future use of survey based SCD when examining cognitive functioning and its risk factors.

Conclusions

At individual level, the HHS score can provide concrete examples and motivation for understanding one's future potential developmental patterns of risk factors and the contributions of their changes to later SCD. It could be used to detect the risk of adverse development early, especially when lifestyle information is included in the risk assessment by the patient or a health professional. Early detection and interventions already in midlife can improve the quality of life and add healthier years to life. From a broader perspective, these findings emphasize the need for AI-driven tools in public health and healthcare policy. Predictive models like the HHS score could be integrated into preventive healthcare initiatives, supporting policymakers in designing evidence-based interventions to reduce the burden of cognitive decline. By utilizing personalized health assessments, healthcare systems can proactively address modifiable risk factors and improve long-term cognitive outcomes in aging populations.

Future research should assess whether long-term lifestyle interventions can reduce dementia and

SCD risk, as demonstrated in the FINGER study [54]. However, showing their effectiveness in preventing clinically diagnosed dementia remains challenging. AI-driven tools could help by modeling long-term intervention effects, integrating biomarkers to improve risk prediction, and supporting personalized prevention strategies. Additionally, the HHS score should be validated in more diverse populations, and other potentially modifiable risk factors of SCD should be further explored.

Acknowledgements

We used ChatGPT, a large language model (LLM) developed by OpenAI to linguistically edit, shorten, and improve the readability of the text. The authors retain full responsibility for the entire content.

Funding

AE and TL are supported by the Research Council of Finland (Grant #330527). AE is supported by the Emil Aaltonen Foundation and Awards - AI in Preventive Health and Social Care - DT4Regions Project Ideathon. AK is supported by the Research Council of Finland Centre of Excellence [Grants #312310, #336669, #352973], OR is supported by the Juho Vainio Foundation (Grant #202500047) and by the Ministry of Education and Culture, Finland (Grant #OKM/76/626/2023). JK is supported by the Research Council of Finland (Grants #355265 and #360990)

Authors' contributions

All authors (AE, TR, MH, IB, JK, OR, AK and TL) contributed to planning the study, critically revised and commented on the manuscript text, and approved the submission of the final version. AE conducted the analyses, with help of TR. AE drafted the first version of the manuscript with help of MH.

Conflict of interest statement

Disclosure		Statement
Financial	Disclosure:	none.
Non-financial	Disclosure:	none

Patient consent

not required.

Ethical approvals

This study was approved and has received research permit from the health authorities of the City of Helsinki. The study plan was approved by the Ethical Committee of the Faculty of Medicine, University of Helsinki, Finland. Permission for the secondary use of health and social data is granted by the Finnish Social and Health Data Permit Authority, Findata, and participants provided informed consent for scientific research. All data were handled in compliance with data protection regulations.

Data availability statement

The Helsinki Health Study survey data cannot be shared due to strict data protection laws and regulations. More information is available at project website[55], in our data protection statement[56], and from kttl-hhs@helsinki.fi. Permission to the secondary use of health and social data is granted by the Finnish Social and Health Data Permit Authority, Findata (info@findata.fi).

Abbreviations

AI: artificial intelligence

BMI: body mass index

CPT: conditional probability table

DBN: dynamic Bayesian networks

HHS: Helsinki Health Study

LTPA: leisure-time physical activity

MAP: maximum a posteriori

qNML: quotient normalized maximum likelihood

SCD: subjective cognitive decline



References

1. World Health Organization. UN Decade of Healthy Ageing: Plan of Action 2021–2030 [Internet]. Geneva, Switzerland: World Health Organization; 2020 [cited 2025 Mar 8]. Available from: <https://cdn.who.int/media/docs/default-source/decade-of-healthy-ageing/decade-proposal-final-apr2020-en.pdf>
2. Scott AJ, Ellison M, Sinclair DA. The economic value of targeting aging. *Nat Aging*. 2021 Jul 5;1(7):616–23.
3. Harper S. Economic and social implications of aging societies. *Science*. 2014 Oct 31;346(6209):587–91.
4. Vollset SE, Ababneh HS, Abate YH, Abbafati C, Abbasgholizadeh R, Abbasian M, et al. Burden of disease scenarios for 204 countries and territories, 2022–2050: a forecasting analysis for the Global Burden of Disease Study 2021. *The Lancet*. 2024 May;403(10440):2204–56.
5. Levine S, Malone E, Lekachvili A, Briss P. Health Care Industry Insights: Why the Use of Preventive Services Is Still Low. *Prev Chronic Dis*. 2019 Mar 14;16:180625.
6. Nyberg ST, Singh-Manoux A, Pentti J, Madsen IEH, Sabia S, Alfredsson L, et al. Association of Healthy Lifestyle With Years Lived Without Major Chronic Diseases. *JAMA Intern Med*. 2020 May 1;180(5):760.
7. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*. 2024 Jul;S0140673624012960.
8. Jessen F, Amariglio RE, Buckley RF, Van Der Flier WM, Han Y, Molinuevo JL, et al. The characterisation of subjective cognitive decline. *The Lancet Neurology*. 2020 Mar;19(3):271–8.
9. Mitchell AJ, Beaumont H, Ferguson D, Yadegarfar M, Stubbs B. Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis. *Acta Psychiatr Scand*. 2014 Dec;130(6):439–51.
10. Baumgart M, Snyder HM, Carrillo MC, Fazio S, Kim H, Johns H. Summary of the evidence on modifiable risk factors for cognitive decline and dementia: A population-based perspective. *Alzheimer's & Dementia*. 2015 Jun;11(6):718–26.
11. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. *The Lancet*. 2017 Dec;390(10113):2673–734.
12. Makmee P, Wongupparaj P. VR Cognitive-based Intervention for Enhancing Cognitive Functions and Well-being in Older Adults with Mild Cognitive Impairment: Behavioral and EEG Evidence. *Psychosocial Intervention*. 2025 Jan 2;34(1):37–51.
13. Dallora AL, Eivazzadeh S, Mendes E, Berglund J, Anderberg P. Machine learning and microsimulation techniques on the prognosis of dementia: A systematic literature review. Chen K, editor. *PLoS ONE*. 2017 Jun 29;12(6):e0179804.

14. Onisko A, Druzdzel MJ, Austin RM. How to interpret the results of medical time series data analysis: Classical statistical approaches versus dynamic Bayesian network modeling. *Journal of Pathology Informatics*. 2016 Jan;7(1):50.
15. Ladyzynski P, Molik M, Foltynski P. Dynamic Bayesian networks for prediction of health status and treatment effect in patients with chronic lymphocytic leukemia. *Sci Rep*. 2022 Feb 2;12(1):1811.
16. Lahelma E, Aittomäki A, Laaksonen M, Lallukka T, Martikainen P, Piha K, et al. Cohort Profile: The Helsinki Health Study. *International Journal of Epidemiology*. 2013 Jun;42(3):722–30.
17. Etholén A, Kouvonen A, Hänninen M, Kulmala J, Rahkonen O, Mänty M, et al. Individual and dual trajectories of insomnia symptoms and body mass index before and after retirement and their associations with changes in subjective cognitive functioning. *Preventive Medicine*. 2023 Dec;107830.
18. Current Care Guidelines [Internet]. 2024. Available from: <http://www.kaypahoito.fi/en/>
19. World Health Organization Regional Office for Europe. Body mass index - BMI. 2021 [cited 2022 Nov 4]; Available from: <https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations>
20. Etholén A, Pietiläinen O, Kouvonen A, Hänninen M, Rahkonen O, Lallukka T. Trajectories of Insomnia Symptoms Among Aging Employees and Their Associations With Memory, Learning Ability, and Concentration After Retirement - A Prospective Cohort Study (2000–2017). *J Aging Health*. 2022 Apr 28;089826432210787.
21. The Finnish Institute for Health and Welfare, (THL). TOIMIA-tietokanta. 2011 Jan 17 [cited 2024 Jun 15]; Available from: <https://www.terveysportti.fi/apps/dtk/tmi/article/tmm00047>
22. Murphy, K. P. *Dynamic bayesian networks: representation, inference and learning*. University of California, Berkeley.; 2002.
23. Murphy K. *Dynamic Bayesian Networks*. 2002 Jan;
24. Pearl J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann; 1988.
25. Microsoft. C# Programming Language [Internet]. 2023. Available from: <https://docs.microsoft.com/en-us/dotnet/csharp/>
26. Jean-Michel Perraud et al. R.NET [Internet]. Available from: <https://www.nuget.org/packages/R.NET/1.9.0>
27. R Core Team. *R: A Language and Environment for Statistical Computing* [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2022. Available from: <https://www.R-project.org/>
28. Scutari M. Learning Bayesian Networks with the **bnlearn** R Package. *J Stat Soft* [Internet]. 2010 [cited 2023 Dec 27];35(3). Available from: <http://www.jstatsoft.org/v35/>

i03/

29. T. Minka, J. Winn, J. Guiver, Y. Zaykov, D. Fabian, J. Bronskill. Infer.NET 0.3 [Internet]. Microsoft Research Cambridge; Available from: <http://dotnet.github.io/infer>
30. Heckerman D. Data Mining and Knowledge Discovery. 2004;1(1):79–119.
31. Schwarz G. Estimating the Dimension of a Model. The Annals of Statistics. 1978;6(2):461–4.
32. Silander, T., Leppä-Aho, J., Jääsaari, E., & Roos, T. Quotient normalized maximum likelihood criterion for learning Bayesian network structures. In A. Storkey, & F. Perez-Cruz (Eds.), International Conference on Artificial Intelligence and Statistics, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands (pp. 948-957). (Proceedings of Machine Learning Research; Vol. 84). JMLR. <http://proceedings.mlr.press/v84/silander18a.html>. In.
33. Aalen O, Røysland K, Gran J, Kouyos R, Lange T. Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms. Stat Methods Med Res. 2016 Oct;25(5):2294–314.
34. Minka TP. Expectation propagation for approximate Bayesian inference. In: Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2001. p. 362–9. (UAI'01).
35. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982 Apr;143(1):29–36.
36. for Cohort Studies of Memory in an International Consortium (COSMIC), Röhr S, Pabst A, Riedel-Heller SG, Jessen F, Turana Y, et al. Estimating prevalence of subjective cognitive decline in and across international cohort studies of aging: a COSMIC study. Alz Res Therapy. 2020 Dec;12(1):167.
37. World Health Organization. Global action plan on the public health response to dementia 2017–2025 [Internet]. Geneva: World Health Organization; 2017 [cited 2024 May 25]. 44 p. Available from: <https://iris.who.int/handle/10665/259615>
38. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. The Lancet. 2015 Jun;385(9984):2255–63.
39. Barha CK, Davis JC, Falck RS, Nagamatsu LS, Liu-Ambrose T. Sex differences in exercise efficacy to improve cognition: A systematic review and meta-analysis of randomized controlled trials in older humans. Front Neuroendocrinol. 2017 Jul;46:71–85.
40. Fondell E, Townsend MK, Unger LD, Okereke OI, Grodstein F, Ascherio A, et al. Physical activity across adulthood and subjective cognitive function in older men. Eur J Epidemiol. 2018 Jan;33(1):79–87.
41. Denche-Zamorano Á, Salas-Gómez D, Franco-García JM, Adsuar JC, Parraca JA,

- Collado-Mateo D. Associations between Physical Activity Frequency in Leisure Time and Subjective Cognitive Limitations in Middle-Aged Spanish Adults: A Cross-Sectional Study. *Healthcare*. 2024 May 22;12(11):1056.
42. Blondell SJ, Hammersley-Mather R, Veerman JL. Does physical activity prevent cognitive decline and dementia?: A systematic review and meta-analysis of longitudinal studies. *BMC Public Health*. 2014 Dec;14(1):510.
 43. Sindi S, Calov E, Fokkens J, Ngandu T, Soininen H, Tuomilehto J, et al. The CAIDE Dementia Risk Score App: The development of an evidence-based mobile application to predict the risk of dementia. *Alz & Dem Diag Ass & Dis Mo*. 2015 Sep;1(3):328–33.
 44. Omura JD, Brown DR, McGuire LC, Taylor CA, Fulton JE, Carlson SA. Cross-sectional association between physical activity level and subjective cognitive decline among US adults aged ≥ 45 years, 2015. *Preventive Medicine*. 2020 Dec;141:106279.
 45. Felez-Nobrega M, Haro JM, Erickson KI, Koyanagi A. Physical Activity is Associated With Fewer Subjective Cognitive Complaints in 47 Low- and Middle-Income Countries. *Journal of the American Medical Directors Association*. 2020 Oct;21(10):1423-1429.e2.
 46. Theou O, Stathokostas L, Roland KP, Jakobi JM, Patterson C, Vandervoort AA, et al. The Effectiveness of Exercise Interventions for the Management of Frailty: A Systematic Review. *Journal of Aging Research*. 2011;2011:1–19.
 47. Stillman CM, Esteban-Cornejo I, Brown B, Bender CM, Erickson KI. Effects of Exercise on Brain and Cognition Across Age Groups and Health States. *Trends in Neurosciences*. 2020 Jul;43(7):533–43.
 48. Knopman D, Boland LL, Mosley T, Howard G, Liao D, Szklo M, et al. Cardiovascular risk factors and cognitive decline in middle-aged adults. *Neurology*. 2001 Jan 9;56(1):42–8.
 49. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative, Franklin JM, Desai RJ, Feldman W, et al. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023 Apr 25;329(16):1376.
 50. Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *The Lancet Neurology*. 2006 Sep;5(9):735–41.
 51. Anstey KJ, Zheng L, Peters R, Kootar S, Barbera M, Stephen R, et al. Dementia Risk Scores and Their Role in the Implementation of Risk Reduction Guidelines. *Front Neurol*. 2022 Jan 4;12:765454.
 52. Thornton JG, Lilford RJ, Johnson N. Decision analysis in medicine. *BMJ*. 1992 Apr 25;304(6834):1099–103.
 53. Anstey KJ, Mack HA, Cherbuin N. Alcohol Consumption as a Risk Factor for Dementia and Cognitive Decline: Meta-Analysis of Prospective Studies. *The American Journal of Geriatric Psychiatry*. 2009 Jul;17(7):542–55.
 54. Lehtisalo J, Rusanen M, Solomon A, Antikainen R, Laatikainen T, Peltonen M, et al.

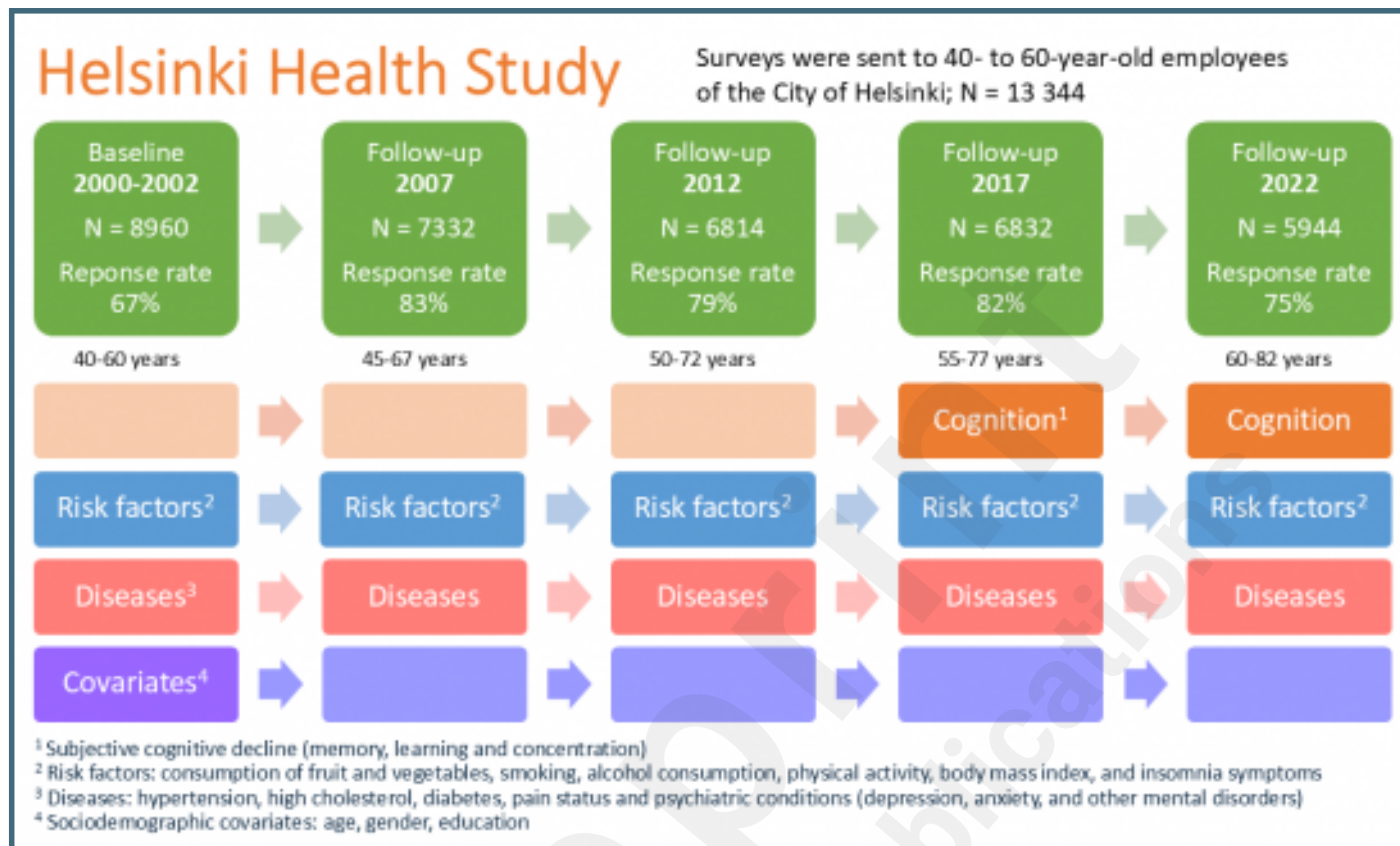
Effect of a multi-domain lifestyle intervention on cardiovascular risk in older people: the FINGER trial. *European Heart Journal*. 2022 Jun 1;43(21):2054–61.

55. Helsinki Health Study - Project website [Internet]. [cited 2024 Jul 30]. Available from: helsinki.fi/hhs
56. Helsinki Health Study - Data protection statement [Internet]. [cited 2024 Jul 30]. Available from: helsinki.fi/hhs_dataprotectionstatement

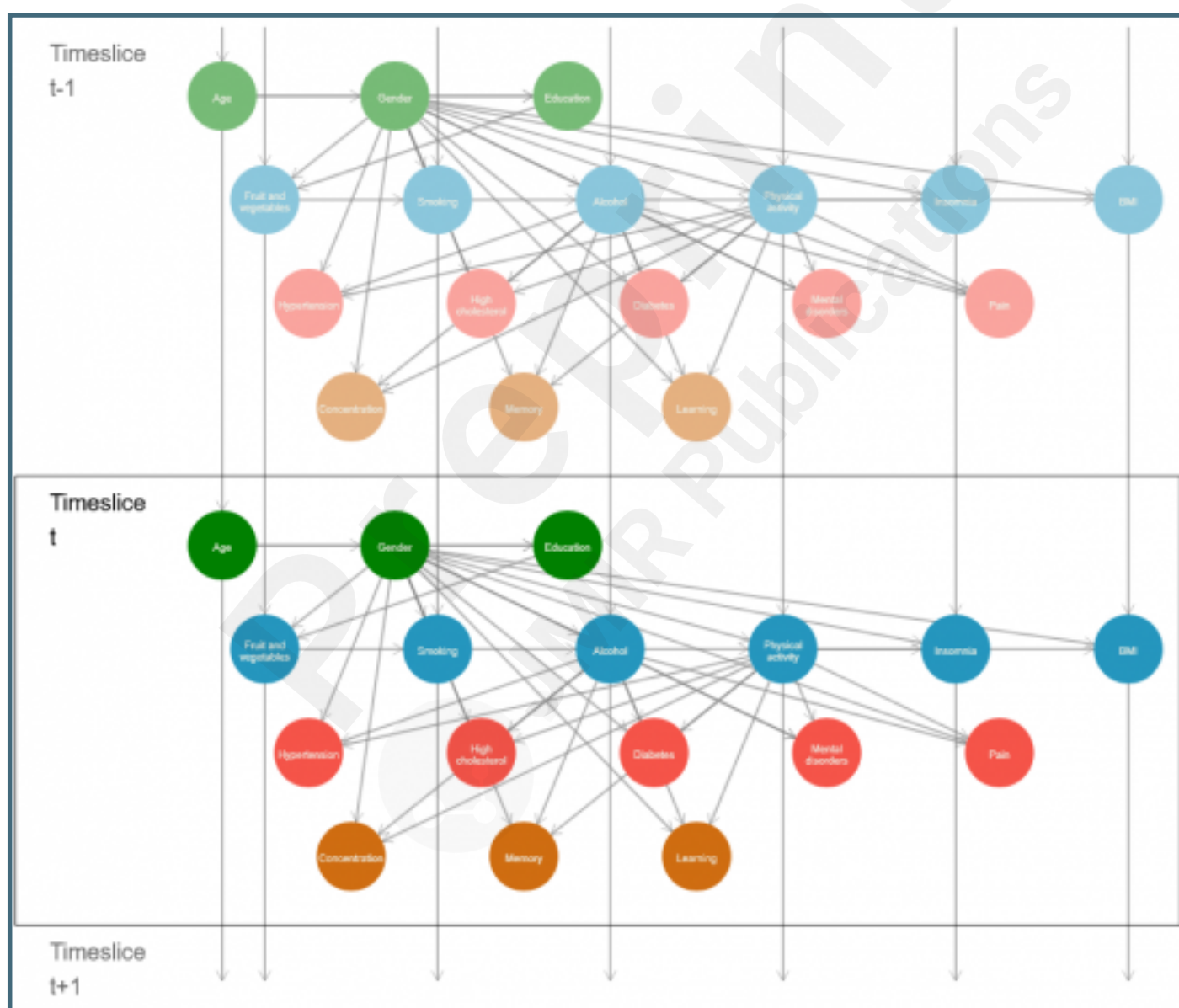
Supplementary Files

Figures

Phases of the Helsinki Health Study cohort: outcome (subjective cognitive decline) and exposures (risk factors) used at different study phases.



The HHS score dynamic Bayesian network model for predicting subjective cognitive decline (SCD) based on the expert knowledge and score-based quotient normalized maximum likelihood criterion algorithm. The flow (possible parents of current node) of network is from left-to-right and from up-to-bottom. Definitions for nodes: Age=Age group (5-year intervals), Gender (woman, man), Education (basic, secondary, higher), Fruit and vegetables = consumption of fruit and vegetables (daily consumption of both, daily consumption of either, non?daily consumption), Smoking (never smoker, ex-smoker, and smoker), Physical activity =leisure-time physical activity (physically inactive: metabolic equivalent (MET) < 14 vs physically active: MET ≥ 14), BMI=body mass index (underweight less than 18.5 or recommend healthy weight 18.5–24.9, overweight 25.0–29.9, obesity 30+), Insomnia=insomnia symptoms (14 nights/month), Hypertension (no, yes), Highcholesterol=high cholesterol(no, yes), Diabetes (no, yes), Mental disorders (no, yes; including depression, anxiety disorder, other mental disorder), Pain=current pain status (no, yes), Concentration (good, declined), Memory (good, declined), Learning (good, declined). Total training size N = 8960 based on five repeated surveys (2000–2022) of the Helsinki Health Study.



The HHS score user interface allows users to their input current risk factors, demographic variables, and potential targets for five years later, generating trajectories of risk factors and risks associated with subjective cognitive decline (SCD) and health factors. For example, in this case the current evidence = 55 years, woman, physically inactive and target evidence = 55 years, woman, physically inactive (55 years), physically active (60 years). The risk of declined memory when 60-year-old based on current evidence = 40.1% and based on target evidence = 33.9% having odds ratio OR 0.76 (95% Bayesian credible interval 0.59–0.99). MET= metabolic equivalent.



The first step in the decision heatmap to decrease self-reported memory decline of a 55-year-old woman with a secondary education. The decision matrix shows effects of possible changes 5 years from now based on the current state. When the first choice is made, the evidence is changed and the new values to matrix are updated making this decision heatmap dynamic.

Current state (now)		Effects of possible changes (or their lack) from current to later self-rated memory decline five years later					
Smoking	Current	Smoking	Alcohol consumption	Physical activity	Fruit and vegetables	Body mass Index	Insomnia
Alcohol consumption	No	Never (no applicable)	No 0%	Inactive 0%	None 0%	Normal -0.5%	None +0.3%
Physical activity	Inactive	Ex-smoker -0.3%	Moderate -10.8%		Either -1.3%	Overweight 0%	Medium +0.1%
Fruit and vegetables	None		High -18.3%				
Body mass index	Overweight	Current 0%	High risk -5.6	Active -15.3%	Both -1.8%	Obesity -0.5%	High 0%
Insomnia	High						

Multimedia Appendixes

Table S1: Conditional probability distribution tables (CPT) to subjective memory decline and leisure-time physical activity in dynamic Bayesian networks.

URL: <http://asset.jmir.pub/assets/d66779955dc10088e032b6a6148de713.xlsx>

Table S2: Spearman rank cross-correlations and autocorrelations between subjective cognitive decline and additional outcome variables in a sample of 35,186 individuals, based on model inference using phased data from our original dataset. Autocorrelations between time slices, $t = 0$ and $t = 1$, are presented on the diagonal, while cross-correlations within $t = 0$ are presented off-diagonal.

URL: <http://asset.jmir.pub/assets/351b9df7d480bc21cd1d183aba087a84.xlsx>

Table S3. Area Under the Receiver Operating Characteristic Curve (AUC) calculated using 5-fold cross-validation for the outcomes of subjective cognitive decline (SCD) and related health conditions of the dynamic Bayesian network (DBN). Total number of participants: 8,960. Time points: $t = 0$ represents the initial step, in which evidence based on gender, age, education, fruit and vegetable consumption, smoking, alcohol consumption, leisure-time physical activity, body mass index, and insomnia symptoms is set in the DBN model; $t = 0$ to $t = 4$ represent outcomes at subsequent time steps, each 5 years apart.

URL: <http://asset.jmir.pub/assets/f740b41f5492b1d3cfa1a356cf7606d5.xlsx>

Methods Appendix: Validation of the subjective cognitive decline (SCD) instrument.

URL: <http://asset.jmir.pub/assets/2b766968d6044ad4ae96824db704e4da.docx>