BMC Geriatrics

RESEARCH

Open Access

Identifying suppressive factors of Alzheimer's disease through comprehensive analysis of real-world data: a single-center retrospective study

Mana Shiotani[1,2], Yuki Hyohdoh[3], Yutaka Hatakeyama[3], Hiroaki Kazui[4] and Yoshiyasu Okuhara[3*]

Abstract

Background In addition to conventional symptomatic treatment drugs, anti-amyloid beta antibody drugs are expected to benefit patients with Alzheimer's disease (AD). However, issues such as side effects and high costs persist, and new preventive and therapeutic drugs are desired. Meanwhile, information on the diagnosis and symptomatic treatment of AD accumulated during daily clinical practice is stored as real-world data and is considered a powerful means of discovering unknown factors that could provide clues for new prevention and treatment approaches for AD through comprehensive exploration.

Methods We used anonymized hospital information system data from a tertiary care and academic hospital in Japan, spanning from 1981 to 2016, to search for potential suppressive factors for AD onset and to verify the validity of the discovered factors. We initially conducted a comprehensive search for candidate suppressive factors for AD and verified them using the inverse probability weighting (IPW) method with propensity scores.

Results From the comprehensive search, we identified glycyrrhizic acid (GA), a component of licorice, a traditional medicine with anti-inflammatory, antioxidant, antibacterial, and antiaging properties, as a candidate suppressing factor for AD. The IPW method showed that the odds ratio of developing AD in the GA group was 0.642 (95% confidence interval: 0.566–0.727) compared with the non-GA group after adjustment.

Conclusions This is the first human study to suggest that GA may be a factor that can suppress the onset of AD. Additionally, our method could be a promising tool for drug repositioning that applies existing drugs already used in clinical settings with well-known side effects to diseases different from their original use.

Keywords Alzheimer's disease, Suppressive factor, Real-world data, Drug repositioning

*Correspondence:
Yoshiyasu Okuhara
okuharay@kochi-u.ac.jp
Full list of author information is available at the end of the article

Background

According to the World Health Organization (WHO), the number of people living with dementia, estimated at 55 million in 2019, is expected to increase to 139 million in 2050 [1]. These patients often present with behavioral and psychological symptoms of dementia (BPSD) in addition to cognitive dysfunction [2]. Patients with BPSD experience emotional distress, diminished quality of life, greater functional impairment, and decreased survival [3]. Caregivers also experience increased burden of stress, depression, and financial consequences, such as decreased income from employment [3].

According to 2021 Alzheimer's disease (AD) facts and figures, AD is the most common cause of dementia, accounting for an estimated 60%–80% of cases [4], and is more commonly observed in individuals aged ≥ 65 years [5]. In Japan, the most aged society in the world, with a population aged ≥ 65 years of 36.21 million, accounting for 28.9% of the total population in 2021 [6], the prevalence of dementia among older adults is expected to rise [7]. This trend is also observed in other countries with aging populations [8]. The social and familial burden of caring for the AD population will be enormous and unsustainable in the near future worldwide [8]. Therefore, preventive methods and fundamental treatments for AD should be established.

Until recently, cholinesterase inhibitors (donepezil, galantamine, and rivastigmine) and N-methyl-D-aspartate receptor antagonist memantine were the only drugs available [9] used as symptomatic treatments for AD. Recently, aducanumab and lecanemab, which are amyloid beta (Aβ)-directed antibody drugs, have become a hot topic. Although the U.S. Food and Drug Administration (FDA) granted accelerated approval to aducanumab in 2021 [10], its development was discontinued in January 2024 [11]. In contrast, lecanemab received traditional approval from the FDA in July 2023 [12], which was also approved in Japan in September 2023 [13]. While lecanemab reduces amyloid plaques in the brain, a defining pathophysiological feature of AD [12], side effects, such as amyloid-related imaging abnormalities (ARIA) with edema and ARIA with cerebral hemorrhages, were more common in the lecanemab treatment group during phase 3 randomized controlled clinical trial, which was evaluated for approval [14].

Furthermore, the estimated annual cost of lecanemab treatment in the United States is approximately 26,500 USD per patient [15, 16]. Thus, effective new preventive or therapeutic drugs are necessary.

Risk factors for dementia include hypertension, diabetes, obesity, dyslipidemia, depression [9, 17], smoking [9, 17–19], and low educational background [9, 18]. Known factors that may suppress AD include leisure activities [9], intake of fish (contains n-3 fatty acid) [9, 18], and regular physical activity [9, 20]. These factors have been confirmed through meta-analyses of randomized controlled trials (RCTs) and observational studies. However, these methods require pre-determined candidate factors, and the number of factors considered is limited, making these methods unsuitable for discovering unknown risk factors or suppressive factors.

Conversely, clinical settings generate real-world data (RWD) through daily diagnosis and symptomatic treatment of AD, which is digitized by hospital information systems (HISs). RWD offers advantages, such as the rapid analysis of large datasets and sufficient statistical power, making it a promising approach for comprehensive exploration that is challenging for RCTs or conventional observational studies. This method may lead to the discovery of new risk factors to elucidate the mechanism of disease onset or to investigate potential suppressors of onset, leading to new prevention and treatment strategies. This study aimed to identify potential suppressors of AD onset through a comprehensive analysis of RWD and to verify the validity of the discovered candidates.

Methods

Data sources

This retrospective observational study was conducted from 1981 to 2016 at Kochi Medical School (KMS) Hospital, a 612-bed tertiary care and academic hospital in Kochi Prefecture, Japan. The ethics committee of KMS approved the study protocol and waived the need for patient informed consent due to the retrospective nature of the study. This study adhered to the Declaration of Helsinki of 1975 (as amended in 1983).

At KMS Hospital, a HIS featuring order entry by physicians has been in use since 1981 [21]. This system collects and stores anonymized patient data, including basic information, visit/admission/discharge dates, laboratory test results, prescriptions, physician-diagnosed disease names, and procedures performed. The name of a disease in the differential diagnosis stage is flagged as "possible" when the disease name is registered in the HIS, and a "confirmed" flag is added when it is confirmed. These data are stored in the data warehouse Retrieval sYstem for Open Medical Analysis-2 (RYOMA2) [22, 23]. All data for this study were obtained from RYOMA2.

Overview of the analysis

We used anonymized HIS data from KMS Hospital, covering the period from 1981 to 2016, to search for potential suppressive factors for AD onset and to verify the validity of the discovered factors. Our analysis comprised two steps: identifying potential suppressive factors

for AD and validating them. Below is an overview of the analysis procedure:

Comprehensive search for "characteristic factors" and identification of "candidate factors"

① AD and non-AD groups were extracted from all data accumulated in RYOMA2 from 1981 to 2016.
② Laboratory test results, prescriptions, and names of concomitant diseases were compared comprehensively between the two groups. Factors showing substantial differences between the two groups were identified as "characteristic factors."
③ We used the presence of AD as the objective variable and performed logistic regression analysis with age and sex, along with the "characteristic factors" identified in ②, as explanatory variables. Odds ratios (ORs) were calculated and adjusted for confounding factors. We evaluated the "characteristic factors" using these adjusted ORs to identify factors that contribute to AD suppression, which are referred to as "candidate factors" hereafter.

Verification of "candidate factors"

④ The "candidate factors" identified in ③ were verified using the propensity score (PS) method.

Details of the analysis method

Extraction of target data

Age, sex, laboratory test results, prescriptions, physician-diagnosed disease names of all patients from 1981 to 2016 were included as initial study group. The initial study group was first divided into a group of patients with a disease name containing the word "Alzheimer's" registered (initial AD group) and a group of patients with no disease name containing the word "Alzheimer's" registered at all (non-AD group). Next, from the initial AD group, patients who were registered only in departments other than psychiatry with a disease name containing the word "Alzheimer's" regardless of whether the disease is possible or confirmed and patients who were registered in psychiatry department with a possible disease name containing the word "Alzheimer's" but were not confirmed after all, were excluded. The patient group that underwent this procedure was designated the "AD group." Furthermore, we excluded those without laboratory tests from the non-AD group.

The reason for using disease names containing the word "Alzheimer's" was to collect as many patients as possible to be included in the AD group, and to prevent patients who should have been in the AD group from being included in the non-AD group if we limited disease names to "Alzheimer's disease" or "dementia of the Alzheimer's type," as patients registered with even a slight difference in name would be included in the non-AD group. Therefore, other than "Alzheimer's disease," the names of diseases in the AD group all contain the words "Alzheimer's" and "dementia," although there are differences in expression, classification by age at diagnosis, and severity.

In Japan, the clinical diagnostic criteria for AD have been based on the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [24] since 1984. Recently, it has been recommended [9] to use the National Institute on Aging-Alzheimer's Association (NIA-AA) [25], which is a revised version of NINCDS-ADRDA. Both diagnostic criteria include the presence of dementia as a key to diagnosis [24, 25], and in many clinical settings, Alzheimer's disease is often thought to mean AD dementia. However, in NIA-AA, the need for diagnostic imaging and biomarkers has also been advocated [25]. For this reason, the AD group in this study, which used data from a university hospital, may include patients in the pre-dementia stage, whose brain pathology was diagnosed with a high probability of being Alzheimer's disease using cerebral blood flow single photon emission computed tomography (SPECT) images and cerebrospinal fluid biomarkers. However, since the pathology is the same, KMS hospital has not made a clear distinction between the registered disease names of Alzheimer's disease and dementia caused by Alzheimer's disease. In addition, it is now understood that there are no histopathological differences due to age [25], and it is thought that it is not necessary to distinguish based on the age of onset. For these reasons, patients in the AD group, defined as the group of patients with a confirmed disease name that includes the word "Alzheimer's," can all be treated as patients with the same disease, although they may be at different stages of progression at the time of diagnosis. In addition, in this study, all disease names that include "Alzheimer's" are collectively referred to as "Alzheimer's disease (AD)."

Comprehensive search for "characteristic factors"

First, we comprehensively searched for items showing significant differences between the AD and non-AD groups. We investigated whether each laboratory test was performed, each drug was prescribed, and each concomitant disease was registered, and calculated the ORs between the two groups for the proportion of patients who underwent laboratory test, were prescribed the drug, and registered the concomitant disease. For example, the OR associated with laboratory test A is {number

of patients in the AD group who underwent laboratory test A/(number of patients in the AD group who did not undergo laboratory test A)}/{number of patients in the non-AD group who underwent laboratory test A/(number of patients in the non-AD group who did not undergo test A)}.

If the laboratory test item, prescribed drug, or concomitant disease name was performed, prescribed, or registered at least once for each patient, it was considered present. For laboratory test items, up to 10 items with ORs significantly >1 (more common in the AD group) at a significance level of 0.05 were extracted in descending order of OR. Similarly, we extracted up to 10 drug items (up to 20 items in total) and up to 20 disease items (up to 40 items in total) with ORs significantly >1 (more common in the AD group) and <1 (less common in the AD group).

Identifying "candidate factors" through logistic regression analysis

The AD group was redefined by adding the condition that either a laboratory test was performed within 31 days before or after the registration of the confirmed AD disease name, or a prescription was issued before the confirmed AD disease name was registered. The non-AD group was redefined as those with history of laboratory tests within 31 days before their last test, or with history of prescriptions throughout their entire period. These were used as the target data for logistic regression analysis. Age and sex were added to the "characteristic factors" selected by the comprehensive search as explanatory variables, and logistic regression analysis was performed with the newly defined AD group or non-AD group as the objective variable (AD group = 1, non-AD group = 0). We excluded items with a variance inflation factor (VIF) of ≥10 from the explanatory variables to prevent multicollinearity.

Age was treated as a continuous variable, which was the age at the date of AD registration in the AD group and the age at the date of the last laboratory test in the non-AD group. Sex was assigned a value of 0 for female and 1 for male. The mean value of each laboratory test was calculated for each patient and converted into a categorical variable. Based on the standard values and various guidelines, items with low and high abnormalities were divided into three categories (the low abnormality category was used as the reference), and items with only high abnormalities or only low abnormalities were divided into two categories (the standard value category or the low abnormality category was used as the reference). Regarding the mean value of each laboratory test, the mean value for the 31 days before and after the date wherein a confirmed disease name, including the word "Alzheimer's" was registered in the psychiatric department was calculated in the AD group, and the mean value for the 31 days immediately preceding the last day of each laboratory test was calculated in the non-AD group.

Additionally, some laboratory test results have missing values. We supposed that this was because the physician did not order certain tests owing to absence of abnormal clinical signs [26]. Therefore, we imputed the missing values by assuming that they were included in the standard value category within the range of standard values. We excluded laboratory test items for which the categorical variable conversion resulted in a category with a sample size of 0 from the explanatory variables to prevent instability of the logistic model. Furthermore, some laboratory test items have the same laboratory test item name but with different laboratory test codes due to changes in the test method. In such cases, because the laboratory test values were considered different, they were treated as separate laboratory tests, converted into categorical variables, and combined into a single laboratory test item.

In the AD group, a score of 1 was assigned if the drug was prescribed in any period before a confirmed disease name, including the word "Alzheimer's," and registered in the psychiatric department; otherwise, it was assigned a value of 0. Conversely, in the non-AD group, a score of 1 was assigned if the drug was prescribed at least once in the patient records; otherwise, it was assigned a value of 0. The same drugs with different dosages or specifications were combined into a single explanatory variable. We excluded drugs for treating AD from the explanatory variables. The presence or absence of a prescription as an explanatory variable for the AD group was determined based on whether it was present before AD diagnosis, and its impact on the onset of AD was considered. Therefore, with some exceptions, such as forgetting to register the disease name, drugs used to treat AD (donepezil, galantamine, rivastigmine, and memantine) should not be prescribed before the AD diagnosis. However, in the Japanese health insurance system, the diagnosis of a disease can be registered within the same month as the time of treatment for the disease, such as prescription, and even if the medication is actually prescribed after the diagnosis, the name of the disease may be registered in the HIS after the prescription, making it appear as if the medication was prescribed before the diagnosis. If these data are included, the causal relationship may be reversed, leading to the false conclusion that patients prescribed donepezil, galantamine, rivastigmine, and memantine are more likely to develop AD. Galantamine, rivastigmine, and memantine are AD-specific medications and will not be prescribed unless there is an AD diagnosis, but donepezil may also be prescribed for dementia with Lewy bodies. For this reason, if a patient is diagnosed with dementia

with Lewy bodies before being diagnosed with AD, there is a possibility that donepezil may have been prescribed before the diagnosis of AD. However, while AD accounts for 60–80% of all dementia cases, dementia with Lewy bodies alone is said to account for around 5% [4, 27]. Therefore, the annual incidence rate of developing AD after being diagnosed with dementia with Lewy bodies is thought to be even lower, and the impact on the results is expected to be small.

For concomitant diseases, in the AD group, a score of 1 was assigned if the disease name was registered in all periods before a confirmed disease name, including the word "Alzheimer's" was registered in the psychiatric department; otherwise, the value was assigned as 0. In the non-AD group, a score of 1 was assigned if the disease name was registered in the patient records; otherwise, the value was assigned as 0. However, for concomitant diseases with the same disease name but different disease name codes were combined into a single explanatory variable.

Using the above conditions, we performed a logistic regression analysis to evaluate which of the laboratory tests, prescribed medications, and concomitant diseases could be risk factors or suppressors of AD onset. Among these, factors that are onset suppressors and can be controlled, such as intervention, were designated "candidate factors" and their validity was examined.

Verification of "candidate factors" using propensity score method

We used the PS method to verify "candidate factors." Differences between groups with and without the "candidate factors" were determined using the same method as in the search for "characteristic factors," and items with large differences were considered "background factors" that determine the presence of a "candidate factors." The same group as the target group in the search for "characteristic factors" was used as the initial cohort, and the group was divided into a "group with candidate factor" exposed to the "candidate factors" and a "group without candidate factor" not exposed to the "candidate factors." Furthermore, patients in the "group with candidate factor" exposed to the "candidate factors" after the date on which confirmed disease name including the word "Alzheimer's" was registered were redefined as belonging to the "group without candidate factor." We excluded those in the "group without candidate factor" without laboratory tests.

A logistic regression model was constructed to calculate the PSs using the presence or absence of a "candidate factors" as the objective variable (presence = 1, absence = 0) and the "background factors," obtained above, age, and sex, as explanatory variables.

When the explanatory variable values corresponding to each patient data were substituted into the logistic regression model equation, the probability of the objective variable being 1 was defined as the PS for each patient [28, 29].

The proportion of patients with AD was calculated by weighting the groups with and without "candidate factors" with 1/PS and 1/(1-PS), respectively, based on the inverse probability weighting (IPW) method [29, 30].

The OR and its confidence interval (CI) of AD onset in both groups were calculated from the ratios calculated using the IPW method. Using this method, the impact of the "candidate factors" on AD onset was evaluated using the ORs adjusted with PSs.

The comprehensive search for "characteristic factors" was performed using Structured Query Language (SQL) on the RYOMA2 relational database. All other analyses were performed in R4.1.3 [31]. Identification of "candidate factors" by logistic regression analysis and PSs by logistic regression models were performed using the "glm" function, and covariate balancing was performed using the "WeightIt" [32] and "cobalt" packages [33]. Statistical significance was set at $p < 0.05$.

Results

Target data

Based on the extraction conditions shown in the methods section, the subjects of the comprehensive search were 649 and 205,831 patients in the AD and non-AD groups, respectively, with a total of 206,480 patients (Fig. 1). The mean age of the AD group was 74 years (range: 42 to 95 years), whereas that of the non-AD group was 48 years (range: 0 to 106 years). The age of the AD group was the age at the time when the name of a disease including the word "Alzheimer's" was first registered in the psychiatric department, and the age of the non-AD group was the age at the time when the last laboratory test was performed.

Comprehensive search for "candidate factors"

Based on the comprehensive search for items that differ between the two groups, we obtained data on laboratory test items, prescribed medications, and concomitant disease names, which are summarized in Table 1.

Identifying "candidate factors" using logistic regression analysis

The data for logistic regression redefined under the conditions shown in the methods section were 194,004 patients, including 458 and 193,546 in the AD and non-AD groups, respectively (Fig. 1). For this data, each laboratory test item obtained based on the comprehensive search was converted into a categorical variable within

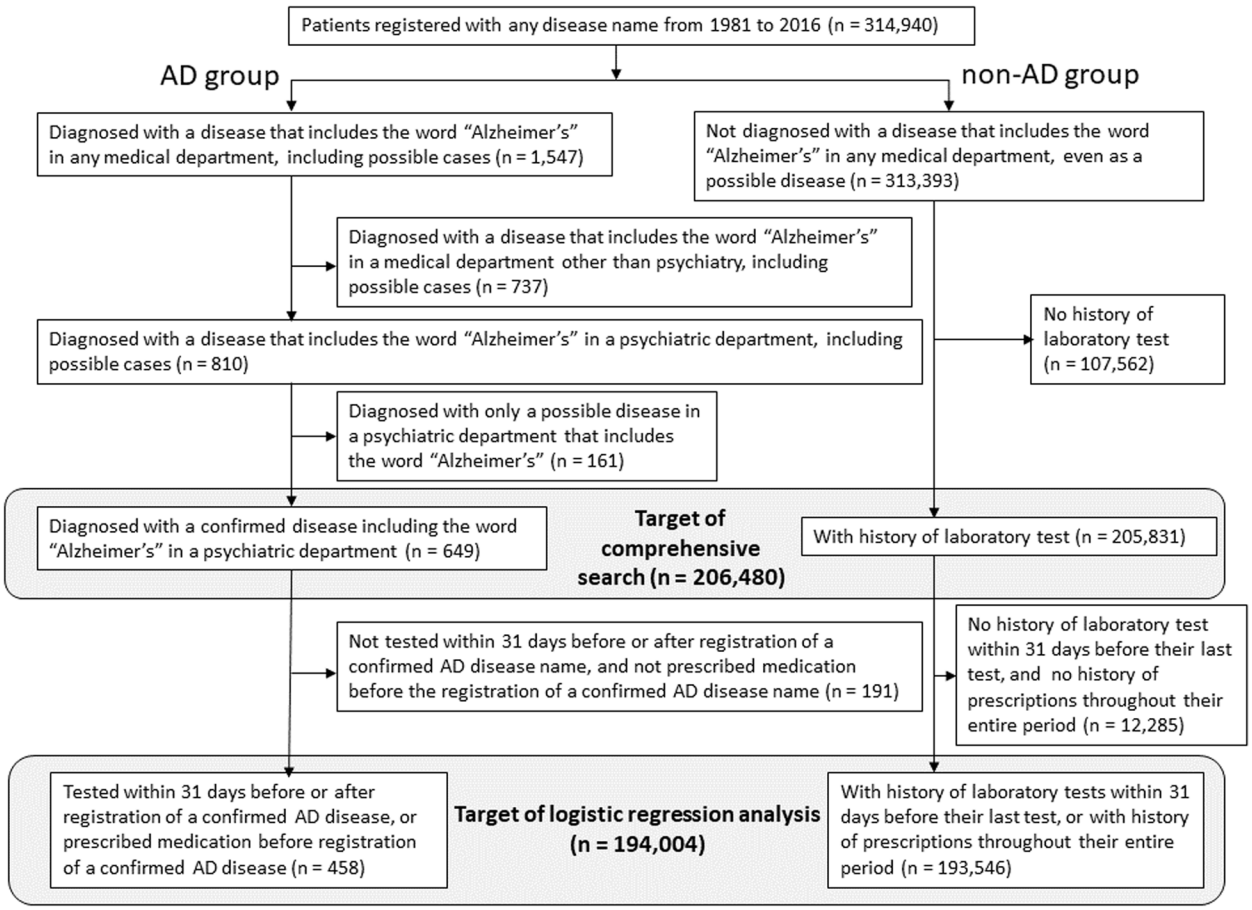

Fig. 1 Flow chart of the comprehensive search and logistic regression analysis

the range of laboratory test values shown in Table 2. In addition, the same drugs or the same disease name with different drug/disease codes, carbazochrome sodium sulfonate hydrate, glucose solution, saline, neurosis, schizophrenia, sleep disturbance, cerebral infarction, and depression were combined into single explanatory variable each. Table 3 shows the descriptive statistics after categorizing the variables and combining the drug and disease names. Category 1 of vitamin B1 (VB1) will be represented as VB1_1. In the case of folate, the reference standard substance used in the folate test at KMS hospital was changed in 2013 from that of the WHO to that of the United States Pharmacopeia, and the reference value was also changed from > 3 ng/mL to ≥ 4 ng/mL [34]. Therefore, the reference value of "Folate①" performed before the change and the reference value of "Folate②" performed after the change were different, even though they had the same test name. For this reason, the test value of each test was converted into a categorical variable based on the respective reference value, and then both tests were treated as one test called " Folate." For example, if "Folate①" was in category 1 or "Folate②" was in category 1, both were ultimately treated as having " Folate" in category 1. There was no overlap between the implementation periods of "Folate①" and "Folate②," but since the average of the test values for a certain period was calculated, if the reference standard substance was changed during this period, both results, "Folate①" and "Folate②," may exist for the same patient. In such cases, for folate, values greater than the cutoff value were considered to be within the reference value, so if either value was smaller than the cutoff value, the final folate value was treated as being in the category smaller than the reference value.

After categorizing the variables, among all the characteristic factors extracted in the previous stage, laboratory test items including categories with a sample size of zero when the test values were converted into categorical variables, which was the case for triiodothyronine (T3) and thyroxine (T4), were excluded. Next, we excluded factors with VIF ≥ 10, which was the case for thyroid-stimulating hormone (TSH). In addition, we excluded donepezil

Table 1 Results of comprehensive search

Characteristic factors
Laboratory test items
Vitamin B1 (VB1)
Vitamin B12 (VB12)
Folate①
Folate②
Free triiodothyronine (free T3)
Brain natriuretic peptide (BNP)
Thyroid-stimulating hormone (TSH)
Free thyroxine (free T4)
Triiodothyronine (T3)
Thyroxine (T4)
Medications
Donepezil hydrochloride tablet 3 and 5 mg
Donepezil hydrochloride orally disintegrating tablet 3 and 5 mg
Tipepidine hibenzate tablet 10 mg
Aniracetam 200 mg
Rivastigmine 4.5, 9, and 13.5 mg
Glycerin 500 mL
Memantine hydrochloride 5 mg
Quetiapine fumarate 50%
Carbazochrome sodium sulfonate hydrate 50 and 100 mg
Glucose solution (5%) 20 and 500 mL
Saline 500 and 1,000 mL
Glycyrrhizic acid 20 mL
Tranexamic acid (10%) 10 mL
Concomitant diseases
Depression①②
Common cold
Neurosis①②
Sleep disturbance①②③
Schizophrenia①②
Dementia
Cerebral infarction①②
Cerebral infarction sequelae
Senile dementia
Angina pectoris
Hypothyroidism
Hypercholesterolemia
Heart failure
Cerebrovascular accident
Cerebral arteriosclerosis

Test items: Tests with the same test name but different laboratory test codes are distinguished by the last number, concomitant disease names: diseases with the same disease name but different disease codes are distinguished based on the last number

hydrochloride, rivastigmine, and memantine hydrochloride as treatments for AD. Aniracetam was used in Japan for improving emotional disorders (anxiety, agitation, and depressed mood) associated with the sequelae of cerebral infarction [35]. Although it may have been used for AD, particularly to alleviate its symptoms, it is not a specific treatment for AD, so it was included in the analysis. Quetiapine fumarate is sometimes used for BPSD, but we did not exclude from the explanatory variables because it is used to treat schizophrenia. The remaining

Table 2 Range of test values for each category

Laboratory test items	Category	Range of test value
VB1 (ng/mL)	1	< 21.3
	2	≥ 21.3, ≤ 81.9
	3	> 81.9
VB12 (pg/mL)	1	< 233
	2	≥ 233, ≤ 914
	3	> 914
Folate① (ng/mL)	1	≤ 3
	2	> 3
Folate② (ng/mL)	1	< 4
	2	≥ 4
Free T3 (pg/mL)	1	< 2.1
	2	≥ 2.1, ≤ 4.1
	3	> 4.1
BNP (pg/mL)	2	≤ 40
	3	> 40
TSH (μU/mL)	1	< 0.34
	2	≥ 0.34, ≤ 3.5
	3	> 3.5
Free T4 (ng/dL)	1	< 0.7
	2	≥ 0.7, ≤ 1.7
	3	> 1.7
T3 (ng/dL)	1	< 80
	2	≥ 80, ≤ 180
	3	> 180
T4 (μg/dL)	1	< 5
	2	≥ 5, ≤ 12
	3	> 12

Range of laboratory test values used when categorizing each laboratory test item. Folate had the same laboratory test name but different laboratory test methods, which were distinguished using the last number

factors, sex and age were also used as explanatory variables in the logistic regression analysis. Table 4 shows the results of the logistic regression analysis. The sensitivity, specificity, accuracy, and Area Under the Curve (AUC) of the logistic regression model were 0.819, 0.784, 0.784, 0.895 respectively. The values of sensitivity, specificity and accuracy correspond to the cutoff value, which was the minimum distance point from the coordinate (1-specificity = 0, sensitivity = 1) in the upper left corner of the Receiver Operating Characteristic curve. The average values of sensitivity, specificity, accuracy, and AUC by tenfold cross validation were 0.810, 0.802, 0.802, 0.885 respectively. The minimum values and maximum values of them were (0.698, 0.905), (0.764, 0.846), (0.764, 0.846), (0.838, 0.929) respectively.

Next, in order to discover factors that could be used for prevention or treatment, factors with significant ORs < 1 that have not been proven to have an effect on suppressing the onset of AD and that can be controlled

Table 3 Descriptive statistics for each variable

Variables	Descriptive statistics	
	AD group (%)	Non-AD group (%)
Age median (Q1–Q3)	76.0 (68.3–81.0)	55.0 (30.0–71.0)
Male	152 (33.2)	92,543 (47.8)
Female	306 (66.8)	101,003 (52.2)
VB1_1	10 (2.18)	696 (0.360)
VB1_2	446 (97.4)	192,509 (99.5)
VB1_3	2 (0.437)	341 (0.176)
VB12_1	9 (1.97)	621 (0.321)
VB12_2	433 (94.5)	190,881 (98.6)
VB12_3	16 (3.49)	2,044 (1.06)
Folate_1	16 (3.49)	1,626 (0.840)
Folate_2	442 (96.5)	191,920 (99.2)
Free T3_1	17 (3.71)	3,106 (1.60)
Free T3_2	431 (94.1)	188,816 (97.6)
Free T3_3	10 (2.18)	1,624 (0.839)
BNP_2	414 (90.4)	187,605 (96.9)
BNP_3	44 (9.61)	5,941 (3.07)
TSH_1	3 (0.655)	1,337 (0.691)
TSH_2	427 (93.2)	184,776 (95.5)
TSH_3	28 (6.11)	7,433 (3.84)
Free T4_1	2 (0.437)	308 (0.159)
Free T4_2	451 (98.5)	192,084 (99.2)
Free T4_3	5 (1.09)	1,154 (0.596)
T3_1	2 (0.437)	335 (0.173)
T3_2	456 (99.6)	193,071 (99.8)
T3_3	0 (0.00)	140 (0.0723)
T4_1	0 (0.00)	150 (0.0775)
T4_2	458 (100)	193,288 (99.9)
T4_3	0 (0.00)	108 (0.0558)
Carbazochrome sodium sulfonate hydrate	57 (12.4)	32,597 (16.8)
Donepezil hydrochloride	49 (10.7)	158 (0.0816)
Tipepidine hibenzate	8 (1.75)	10,315 (5.33)
Aniracetam	52 (11.4)	426 (0.220)
Rivastigmine	10 (2.18)	6 (0.00310)
Glucose solution	31 (6.77)	21,660 (11.2)
Glycerin	13 (2.84)	10,363 (5.35)
Saline	41 (8.95)	24,786 (12.8)
Memantine hydrochloride	5 (1.09)	3 (0.00155)
Quetiapine fumarate	20 (4.37)	375 (0.194)
Glycyrrhizic acid	31 (6.77)	18,161 (9.38)
Tranexamic acid	46 (10.0)	25,670 (13.3)
Senile dementia	21 (4.59)	184 (0.0951)
Dementia	46 (10.0)	513 (0.265)
Cerebral infarction sequelae	37 (8.08)	1,444 (0.746)
Neurosis	82 (17.9)	3,112 (1.61)
Schizophrenia	68 (14.8)	2,829 (1.46)
Sleep disturbance	53 (11.6)	6,632 (3.43)
Depression	57 (12.4)	5,560 (2.87)

Table 3 (continued)

Variables	Descriptive statistics	
	AD group (%)	Non-AD group (%)
Cerebral infarction	107 (23.4)	9,103 (4.70)
Cerebral arteriosclerosis	11 (2.40)	1,794 (0.927)
Cerebrovascular accident	50 (10.9)	4,249 (2.20)
Heart failure	74 (16.2)	5,988 (3.09)
Hypercholesterolemia	26 (5.68)	3,366 (1.74)
Hypothyroidism	82 (17.9)	6,728 (3.48)
Angina pectoris	38 (8.30)	5,158 (2.67)
Common cold	18 (3.93)	12,020 (6.21)

Category 1 of VB1 is represented as VB1_1

(e.g., intervention) were designated as "candidate factors." Logistic regression analysis showed that the ORs were significantly < 1 for male, glucose solution, glycirrhizic acid (GA), and common cold, of which glucose solution and GA were considered to be interventions. In Japan, GA was widely used as a symptomatic treatment for chronic hepatitis before the establishment of a cure for viral hepatitis C, and is still used today to treat not only hepatitis but also eczema, dermatitis, urticaria, pruritus, stomatitis, phlycten, drug rash, and toxic rash. It is known that GA has an anti-inflammatory effect [36]. Also, it has been demonstrated that GA selectively inhibits the activities of arachidonate cascade related enzymes [36]. On the other hand, many studies have pointed out the connection between AD and inflammation in the brain [37, 38]. Therefore, we considered the anti-inflammatory effects of GA to be effective in preventing the onset of AD and included it as a "candidate factor". Regarding glucose, which also showed a significant OR < 1, some studies have suggested a relationship between AD and glucose metabolism disorders [39–41]. For this reason, it may be a factor that suppresses the onset of AD, however, since it is used temporarily and is unlikely to have a continuous pharmacological effect, it was not included as a candidate.

Propensity score method

The data subject to the comprehensive search for selecting explanatory variables for the logistic regression model for PS calculation was 206,408 patients after imposing the conditions shown in the methods section. Of these, 20,909 and 185,499 were in the groups with "candidate factor" GA (GA group) and without "candidate factor" GA (non-GA group), respectively (Fig. 2). Although, the "candidate factor" GA found in the comprehensive search was an injectable agent (monoammonium glycyrrhizinate for intravenous injection 20 mL), we also included

Table 4 Results of logistic regression analysis

Variables	OR (95%CI)
Age	1.05 (1.05–1.06)***
Male	0.64 (0.52–0.78)***
VB1_2	0.53 (0.26–1.08)
VB1_3	0.22 (0.04–1.14)
VB12_2	0.48 (0.23–1.03)
VB12_3	0.45 (0.18–1.12)
Folate_2	0.75 (0.42–1.34)
Free T3_2	1.47 (0.86–2.52)
Free T3_3	2.14 (0.90–5.12)
BNP_3	0.68 (0.47–1.00)
Free T4_2	0.38 (0.09–1.61)
Free T4_3	0.48 (0.09–2.60)
Carbazochrome sodium sulfonate hydrate	0.66 (0.40–1.08)
Tipepidine hibenzate	0.91 (0.44–1.90)
Aniracetam	18.27 (12.29–27.15)***
Glucose solution	0.55 (0.37–0.83)**
Glycerin	0.91 (0.50–1.67)
Saline	0.71 (0.49–1.04)
Quetiapine fumarate	2.1 (1.17–3.77)*
Glycyrrhizic acid	0.61 (0.40–0.92)*
Tranexamic acid	0.92 (0.53–1.58)
Senile dementia	11.65 (6.88–19.73)***
Dementia	5.80 (3.93–8.56)***
Cerebral infarction sequelae	1.77 (1.16–2.71)**
Neurosis	3.27 (2.35–4.55)***
Schizophrenia	2.82 (1.96–4.07)***
Sleep disturbance	0.79 (0.55–1.14)
Depression	1.81 (1.30–2.53)***
Cerebral infarction	1.63 (1.24–2.15)***
Cerebral arteriosclerosis	0.88 (0.46–1.68)
Cerebrovascular accident	1.63 (1.14–2.34)**
Heart failure	2.18 (1.57–3.02)***
Hypercholesterolemia	1.81 (1.16–2.82)**
Hypothyroidism	2.98 (2.20–4.04)***
Angina pectoris	1.15 (0.79–1.68)
Common cold	0.53 (0.31–0.91)*

OR Odds ratio, *CI* Confidence interval

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

oral agents of glycyrrhizic acid as GA. Table 5 shows the items selected based on the comprehensive comparison between the GA and non-GA groups. A logistic regression model was constructed with these items, with age and sex being explanatory variables, and whether the patient was in the GA (=1) or non-GA group (=0) being the objective variable. The sensitivity, specificity, accuracy, and AUC of the model were 0.663, 0.761, 0.751,

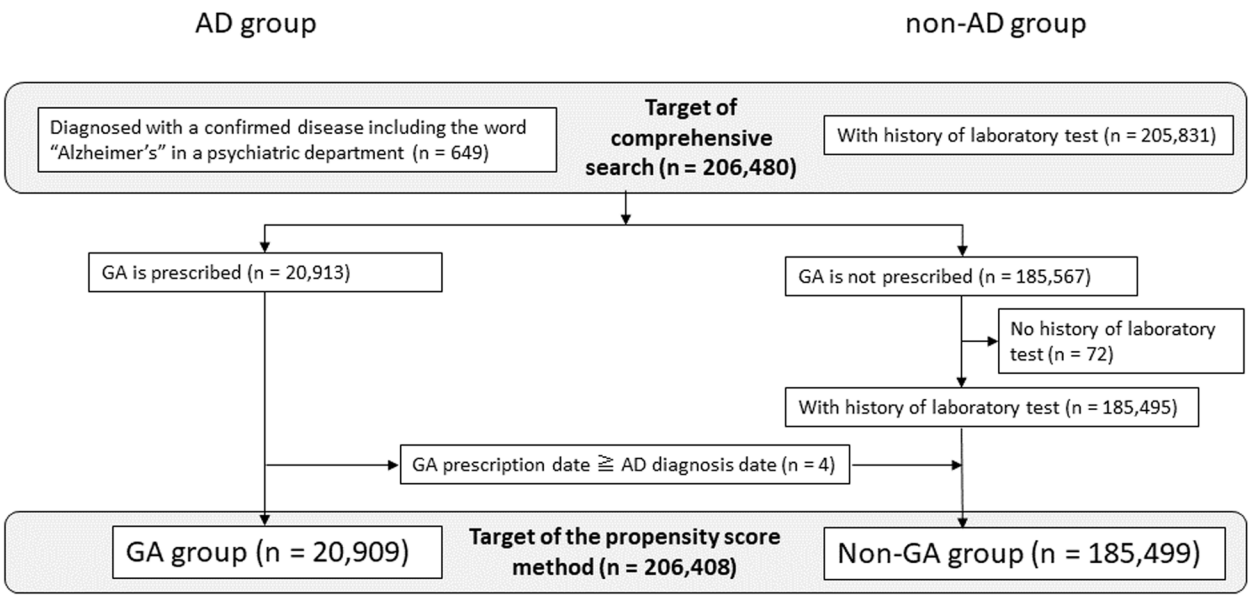

Fig. 2 Flow chart of data extraction using the propensity score method

0.764 respectively. The average values of sensitivity, specificity, accuracy, and AUC by tenfold cross validation were 0.668, 0.756, 0.747, 0.763 respectively. The minimum values and maximum values of them were (0.645, 0.691), (0.734, 0.777), (0.728, 0.764), (0.755, 0.781) respectively. Figure 3 shows that all the covariates adjusted by the PSs calculated using this model had adjusted absolute mean differences of ≤ 0.1. Using these PSs, we used the IPW method to calculate the OR of developing AD in the GA group compared with the non-GA group. The estimated OR before adjustment was 0.660 (95% CI: 0.479–0.911, $p = 0.011$), and the estimated OR after adjustment was 0.642 (95% CI: 0.566–0.727, $p < 0.001$).

Discussion

Geographic scope of sample selection

KMS Hospital is the only university hospital in Kochi Prefecture, where 0.5% of Japan's population resides (2022) [42], and travel time to the hospital, located almost in the center of the prefecture, is approximately less than 3 h by car from the entire prefecture [43]. For this reason, it is considered that the hospital covers patients within the prefecture without geographical bias. According to statistics on discharged patients in 2022, the proportion of patients from outside the prefecture is 1.2%, and most patients live within the prefecture. There are no problems with travel between Kochi Prefecture and other regions via highways, rail ways and air routes, and it is unlikely that geographical conditions will make a significant difference to the sample population compared to other regions of Japan.

Identifying "candidate factors" using logistic regression analysis

Logistic regression model

The sensitivity, specificity, accuracy, and AUC of the logistic regression model used in identifying "candidate factors" and the evaluation with cross validation showed validity of the model.

Significant factors with odds ratios > 1

Older age [5] and hypercholesterolemia [9, 17] were consistent with or indicated in previous studies. As mentioned in the results section, aniracetam may have been used to alleviate AD symptoms, and quetiapine fumarate has been used for BPSD. Differential diagnoses may include senile dementia, dementia, schizophrenia, depression, hypothyroidism, cerebral infarction sequelae, cerebral infarction, and cerebrovascular accident [9, 44]. Heart failure (HF) and AD are known to be associated [45]. Cerebrovascular disease may contribute to neurodegeneration [46], which may be why the ORs associated with cerebral infarction sequelae, cerebral infarction, and cerebrovascular accidents were significantly >1. As for neurosis, its main symptoms are anxiety and mild depression, and these symptoms are also frequently observed in people with early-stage dementia mostly caused by AD [47]; thus, it is believed that this disease name was often added as differential diagnosis. These results were consistent with or interpretable in previous studies, demonstrating the validity of our methods.

Table 5 Variables included in the model for calculating propensity scores

Item name	Abbreviation in Fig. 3	Range of test value
Laboratory test items		
Indocyanine green (ICG) retention rate at 15 min (%)	ICG_1	≤ 10
	ICG_2	> 10
Arterial blood buffer base (ABB) (mEq/L)	ABB_1	< 45.5
	ABB_2	≥ 45.5, ≤ 50.5
	ABB_3	> 50.5
Phenolsulfonphthalein (PSP) excretion rate at 120 min (%)	PSP120_1	< 55
	PSP120_2	≥ 55, ≤ 85
	PSP120_3	> 85
Prothrombin time (PT) control value	PTC_1	< 9.5
	PTC_2	≥ 9.5, ≤ 12.4
	PTC_3	> 12.4
Activated partial thromboplastin time (APTT) control value	APTTC_1	< 26.7
	APTTC_2	≥ 26.7, ≤ 34.0
	APTTC_3	> 34.0
α-Fetoprotein (AFP) (ng/mL)	AFP_1	≤ 10
	AFP_2	> 10
Fibrin and fibrinogen degradation products (FDP) semiquantitative analysis (μg/mL)	FDP_1	≤ 10
	FDP_2	> 10
PT (s)	PPT_1	< 10
	PPT_2	≥ 10, ≤ 13.5
	PPT_3	> 13.5
APTT (s)	APTT_1	< 20
	APTT_2	≥ 20, ≤ 40
	APTT_3	> 40
Immunosuppressive acidic protein (IAP) (μg/mL)	IAP_1	< 202
	IAP_2	≥ 202, ≤ 451
	IAP_3	> 451
Medications		
Halothane	Halothane	
Indocyanine green	Indocyanine green	
Pancuronium bromide	Pancuronium	
Cellulose, Oxidized	Cellulose_Oxidized	
Amidotrizoic Acid	Amidotrizoic	
Menatetrenone	Menatetrenone	
Neostigmine methylsulfate	Neostigmine	
Hypotonic electrolyte infusion (maintenance fluid) with 10% dextrose	Maintenance fluid	
Buprenorphine hydrochloride	Buprenorphine	
Vitamin mixture for high calorie parenteral nutrition infusion	Vitamin_nutrition infusion	
Cefdinir	Cefdinir	
Cefazolin sodium hydrate	Cefazolin	
Betamethasone sodium phosphate	Betamethasone	
Remifentanil hydrochloride	Remifentanil	
Cefcapene pivoxil hydrochloride hydrate	Cefcapene	
Loxoprofen sodium hydrate	Loxoprofen	
Acetaminophen	Acetaminophen	
Rebamipide	Rebamipide	
Fentanyl citrate	Fentanyl	
Povidone iodine	Povidone iodine	

Table 5 (continued)

Item name	Abbreviation in Fig. 3	Range of test value
Concomitant diseases		
Abnormal hepatic function	Abnormal hepatic function	
Chronic hepatitis	Chronic hepatitis	
Diabetes mellitus	Diabetes	
Low back pain	Low back pain	
Hypertension	Hypertension	
Constipation	Constipation	
Angina pectoris	Angina pectoris	
Chronic gastritis	Chronic gastritis	
Eczema	Eczema	
Cervical erosion	Cervical erosion	
Astigmatism	Astigmatism	
Common Cold	Common cold	
Pneumonia	Pneumonia	
Gastric ulcer	Gastric ulcer	

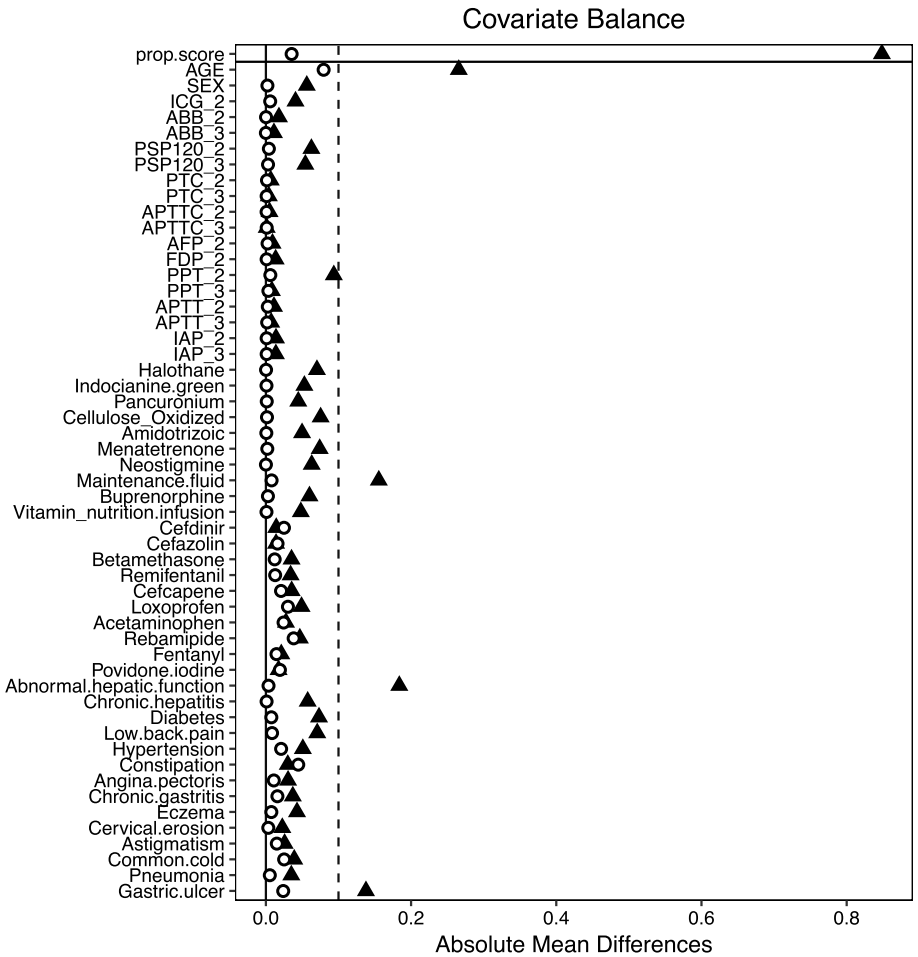

Fig. 3 Covariate Balance ▲: Unadjusted, ○: Adjusted

Factors with significant odds ratios < 1

The significantly lower prevalence in men is consistent with previous studies [4]. The effects of glucose solution and common cold could not be interpreted, however, glucose is often administered transiently and unlikely to be related to AD onset, so we excluded it as a "candidate factor." Although previous studies have suggested the possibility of GA preventing AD onset, no studies have targeted humans, so we selected it as the "candidate factor."

Factors with odds ratios close to 1 and not significant

As is shown in Table 3 and Table 4, despite the bias between AD group and non-AD group, explanatory variables for which the ORs were close to 1 as the result of logistic regression analysis included tipepidine hibenzate, tranexamic acid, glycerin, cerebral arteriosclerosis, and angina pectoris, which had adjusted ORs in the range of 0.85 to 1.15. Tipepidine hibenzate is a drug used to treat coughing and difficulty in expectorating phlegm associated with common cold, and tranexamic acid is a drug used to treat symptoms associated with pharyngitis such as sore throat, redness, congestion, and swelling. They are both thought to have been used to treat the common cold. Although they were used less frequently in the AD group, the ORs from the logistic regression analysis was close to 1 and was not significant. On the other hand, the logistic regression analysis showed that the common cold was significantly less common in the AD group, suggesting that the common cold was a confounding factor for tipepidine hibenzate and tranexamic acid. Cerebral arteriosclerosis was more common in the AD group, but cerebral infarction, cerebrovascular accident, and cerebral infarction sequelae may have been confounding factors. Angina pectoris was also more common in the AD group, possibly because the AD group was more elderly. In the logistic regression analysis including age as a covariate factor, the OR associated with angina pectoris was not significant. This suggests that angina pectoris was not an independent factor. The results for glycerin could not be interpreted.

Relationship between laboratory test variables and Alzheimer's disease, biological mechanisms and clinical significance

Vitamin B1

VB1 deficiency has long been shown to be associated with neurological problems such as cognitive impairment and encephalopathy [39]. In addition, there are many similarities between classical VB1 deficiency and AD [39]. Both are accompanied by cognitive impairment and impaired glucose metabolism in the brain [39]. VB1-dependent enzymes are key components of glucose metabolism that are reduced in the brains of AD patients and by VB1 deficiency, and their reduction may be responsible for the impaired glucose metabolism [39]. Although there are also reports showing that serum VB1 is significantly reduced in AD patients, the analysis method used was analysis of covariance using aligned rank transformation, and the only covariate adjusted was age [48], they are different from those of this study.

Vitamin B12

It has been reported that the serum vitamin B12 (VB12) concentration in patients with AD is significantly lower than that in controls, and the serum homocysteine (Hcy) concentration is significantly higher than that in controls [49]. According to a review by Robinson et al., a possible mechanism for the relationship between VB12 or folate and AD is that reduced folate reduces the availability of methyl donors in the methylation cycle, which is exacerbated by lack of VB12 [50]. This leads to Hcy accumulation and decreased methylation [50]. Promoter demethylation of presenilin-1 and beta-secretase 1 leads to Aβ accumulation and increased phosphorylated tau levels, which are causative agents of AD [50]. However, a meta-analysis concluded that the association between VB12 and AD risk was inconclusive [51].

Folate

Accumulating evidence indicates that AD patients have higher Hcy and lower folate and VB12 levels than healthy controls [51]. In addition, it has been reported that elevated plasma Hcy and low serum folate concentrations are independent predictors of the onset of AD [52]. A possible mechanism for the relationship between folate and AD is as mentioned in the section on VB12. The aforementioned study [52] used a proportional hazards regression model and adjusted for age, sex, education, apolipoprotein E genotype, vascular risk factors, and serum concentrations of folate and VB12, which are different analytical methods and covariates from those in the present study.

Free triiodothyronine and free thyroxine

Free T3 and free T4 are laboratory tests mainly performed to diagnose thyroid dysfunction. Hypothyroidism is included as a characteristic factor in the concomitant disease name, and it is thought that these tests were performed to diagnose this disease. However, some studies suggest a relationship between these test values and AD. In vitro studies have shown that T3 is inversely associated with amyloid precursor protein gene expression, and that T3 deficiency may deprive the brain of its neuroprotective properties in neurodegenerative lesions and Aβ-induced glutamate-mediated excitotoxicity [53]. In

addition, a meta-analysis has reported that free T3 values are significantly lower in AD patients than in controls, but no significant difference was observed in free T4 values [54].

Brain natriuretic peptide

Cardiovascular disease (CVD) and CVD risk factors (hypertension, dyslipidemia, diabetes mellitus) have been suggested to be associated with cognitive decline and may contribute to the development of AD [55]. Brain natriuretic peptide (BNP) and N-terminal pro-brain natriuretic peptide (NT-proBNP) are the most important humoral indicators of cardiac function and HF [55]. However, the molecular mechanism by which BNP plays a role in AD-type cognitive impairment is unclear [55]. In the process of searching for candidate factors in this study, HF was extracted as a common concomitant disease name in patients registered with AD. This is thought to be related to the extraction of BNP as a characteristic factor. However, various opinions have been presented regarding the presence or absence of a relationship between BNP and NT-proBNP, the precursor of BNP, and AD [55–59]. The relationship between BNP or NT-proBNP and AD remains controversial, and this study also did not reach significant results for BNP.

As such, all of these tests have been suggested to be related to AD, and it is understandable that they were extracted as "characteristic factors" in this study. All ORs in the logistic regression analysis showed the same tendency with preceding clinical studies or meta-analyses, except for free T3. That is, for VB1, VB12, folate, free T4, and BNP, the ORs of the onset of AD tended to be smaller than 1 as the test values were higher, while free T3 showed the opposite tendency. However, none of the test items showed statistically significant results. Since the test values were converted into categorical variables in this study, it is possible that the statistical power was insufficient. There are also differences in the analysis methods and the covariates considered between the previous study and this study, which are thought to have influenced the differences in the results.

Propensity score method

The sensitivity, specificity, accuracy, and AUC of the logistic regression model used to calculate PS and the evaluation with cross validation show validity of the model. Figure 3 shows that the covariates adjusted by the calculated PSs were well balanced, indicating that bias was well adjusted. Using these PSs, the OR of AD onset in the GA group relative to the non-GA group adjusted using the IPW method was significantly < 1, indicating that GA can suppress AD onset.

Glycyrrhizic acid

Thus, this study indicates that GA may suppress AD onset. GA, a component of licorice, is a traditional medicine with numerous anti-inflammatory, antioxidant, antibacterial, and antiaging properties [60]. GA was approved in Japan at the latest in 1948, and since then have been widely used. Therefore, information on side effects is well known, and there have been no serious safety issues. In addition, the drug price is extremely low, at 122 yen (=0.79 USD as of February 1, 2025) for 20 mL of monoammonium glycyrrhizinate injection [61].

Several in vitro and in vivo studies have reported results regarding the relationship between GA and AD. Mitochondrial dysfunction is a prominent feature of AD [62–64], and Wang have demonstrated that Aβ, the causative agent of AD, causes mitochondrial dysfunction [63, 64]. Zhu has reported that GA promoted the expression of mitochondrial synthesis regulators and improved cognitive impairment in AD mice [65]. Aβ activates microglia in the brain and causes neuronal death by releasing cytokines [66]. Zhao has reported that GA suppressed microglial activation and reduced memory impairment in AD mice [67].

The protein High Mobility Group Box 1 (HMGB1) promotes AD in parallel with Aβ [68]. GA is a pharmacological inhibitor of HMGB1 [60], and Kong has reported that GA suppressed HMGB1 expression in the cytosol and reduced the severity of memory impairment in aged mice [69].

In addition, results from a study by Tabuchi et al. have shown that GA can reach the brain through the blood–brain barrier [70].

Thus, several previous studies have examined the relationship between AD and GA, indicating that GA may be an effective drug for AD, but these studies have been in vitro or in vivo studies on mice, and no studies have yet been conducted on humans.

Strong points of this study

The strong point of our method is that the patients were divided into two groups, cases and controls, which were almost complementary to each other, and all factors that could be analyzed data-wise were compared comprehensively between the AD and non-AD groups, making it possible to discover factors that were not predicted in advance. Usually, when adjusting for confounding factors in retrospective observational studies that cannot randomly assign comparison groups using multivariate analysis such as logistic regression analysis or the PS method, possible confounding factors are selected based on preceding studies or medical knowledge. However, confounding factors only have effects when the ratios of

those factors are biased between the two groups being compared. Therefore, in this study, factors that can be obtained from the data were comprehensively compared between groups that divide large-scale data into two, and those that are significantly biased between the two groups were selected as possible confounders. For this reason, it is believed that factors that may have effects were covered if they were factors included in the structured data of the HIS, such as prescriptions, laboratory test results, and names of concomitant diseases. This method may be a promising tool for drug repositioning [71], which applies existing drugs that have been used in clinical settings and have well-known side effects to diseases completely different from their original use.

Study limitations

The limitations of this study include the following:

1. As mentioned in the definition of the AD group explained in Methods, the AD group in this study included AD with different degrees of progression. Therefore, results may have varied if stratified by the stage of AD.

   However, this study focused on the preventive effects of disease onset rather than the therapeutic effects after disease onset mitigates this issue.
2. The study was conducted at a single tertiary care and academic hospital, which may limit the generalizability of the findings to other settings. For example, we have more patients with severe symptoms for each disease than general hospitals, which may result in bias. The same is true for AD, where there may be more patients with advanced symptoms than in other hospitals, which could be causing some bias.
   Although there are disadvantages such as limited sample size and institutional bias, the single-institution setting ensured quality control of laboratory tests and provided long-term data, facilitating comprehensive comparisons. In the future, we are considering conducting large-scale data analysis using, for example, the National Database (NDB) provided by the Ministry of Health, Labor and Welfare of Japan [72]. NDB contains two elements: 1) medical receipt information issued by medical institutions to medical insurers, and 2) specific health checkup for people aged 40 to 74 [72], includes questionnaires on lifestyle habits, physical measurements, laboratory test results [73], in a form that does not identify individuals. Since Japan has a universal health insurance system, it covers almost all Japanese people and medical institutions, it will be possible to overcome the problems of sample size and sample bias that were limitations in this study.
3. Because of the exclusion maneuver explained in the methods section, there should be a certain number of cases that may be actually included in AD group but were not included in the analysis, the "false negative group." This may have reduced the AD group from the true population and affected the comparison of the two groups in the analysis. In addition, a reduction in sample size should at least reduce the statistical power. However, on the other hand, it prevents patients who are not actually in AD group from being mixed into the AD group ("false positive group"), which is thought to work in the opposite direction to emphasize the difference in the comparison of the two groups. Besides, even if the statistical power is reduced, if a statistically significant result is obtained, it means that a significant result was obtained despite conservative conditions for statistical significance, and it is thought that the results will be trustworthy. It is possible that patients who should have been included in the non-AD group, for example patients who were suspected of having AD but did not actually have AD, were not included in the non-AD group, which may have influenced the results. However, because the non-AD group was very large, this impact is considered to be small.
4. The analysis relied on structured data from HIS and did not consider unstructured data, such as lifestyle-related factors. Lifestyle-related factors might have been hidden confounding factors.
   Although challenging, future research should develop methods to extract and analyze information from the free-text portions of electronic medical records through text mining. Analysis using NDB data which contains lifestyle-related factors will also overcome this problem.
5. The study could not confirm the dose-response relationship of GA with AD due to variability in prescription amounts and periods.
   Analysis using NDB data which contains enormous amount of data will also overcome this problem.
6. It would be desirable to conduct subgroup analyses based on age, sex, etc. to evaluate the impact on the progression of AD; however, since this study used data from a single institution, the size of the AD-developing group was not sufficient.
   Therefore, such subgroup analyses would also be a task for future analyses of NDB analyses.
7. It is difficult to clarify the pharmacological mechanism of the GA on AD by the method used in this study.

To overcome this problem, network pharmacology [74–76], which has attracted attention as a means of comprehensively elucidating the pharmacological effects of herbal medicines for various diseases, may be useful. Using network pharmacology techniques to identify important proteins that mediate the effects of GA on AD is likely to be effective in clarifying the effectiveness of GA on AD.

Conclusions

Although this study has unavoidable limitations related to RWD as mentioned in study limitations, it effectively used structured data and comprehensively compared factors between patients with and without AD. This approach yielded results consistent with findings from previous studies for many factors. A new finding is that GA may be a factor that suppresses AD onset. While our study alone cannot confirm that GA has a preventive or therapeutic effect on AD, it represents the first human study suggesting this possibility and proposes a hypothesis for future investigation. Future research using NDB will be required to thoroughly examine the relationship between GA prescription amount, prescription period, and its effects. This study also demonstrates that comprehensive comparison methods can be effective for discovering new knowledge and may serve as a promising tool for drug repositioning. Furthermore, network pharmacology techniques would be effective in clarifying the pharmacological mechanism of the discovered drugs on target diseases.

Abbreviations

WHO	World Health Organization
BPSD	Behavioral and psychological symptoms
AD	Alzheimer's disease
Aβ	Amyloid beta
FDA	U.S. Food and Drug Administration
ARIA	Amyloid-related imaging abnormalities
RCT	Randomized controlled trial
RWD	Real-world data
HIS	Hospital information system
KMS	Kochi Medical School
RYOMA2	Retrieval sYstem for Open Medical Analysis-2
OR	Odds ratio
PS	Propensity score
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association
NIA-AA	National Institute on Aging-Alzheimer's Association
SPECT	Single Photon Emission Computed Tomography
VIF	Variance inflation factor
IPW	Inverse Probability Weighting
CI	Confidence interval
SQL	Structured Query Language
VB1	Vitamin B1
VB12	Vitamin B12
free T3	Free triiodothyronine
BNP	Brain natriuretic peptide
TSH	Thyroid-stimulating hormone
free T4	Free thyroxine
T3	Triiodothyronine
T4	Thyroxine
AUC	Area Under the Curve
GA	Glycyrrhizic acid
ICG	Indocyanine green
ABB	Arterial blood buffer base
PSP	Phenolsulfonphthalein
PT	Prothrombin time
APTT	Activated partial thromboplastin time
AFP	α-Fetoprotein
FDP	Fibrin and fibrinogen degradation products
IAP	Immunosuppresive acidic protein
HF	Heart failure
Hcy	Homocysteine
CVD	Cardiovascular disease
NT-proBNP	N-terminal pro-brain natriuretic peptide
HMGB1	High Mobility Group Box 1
NDB	National Database

Acknowledgements

The authors would like to thank Enago (www.enago.jp) for the English language review.

Authors' contributions

M.S. designed the study, provided framework of the manuscript, analyzed the data and drafted the manuscript. Yuki Hyohdoh contributed to the design of the study and drafted the manuscript. Yutaka Hatakeyama contributed to the design of the study and drafted the manuscript. H.K. contributed to the interpretation of results. Y.O. designed the study, provided framework of the manuscript, analyzed the data, drafted the manuscript. All authors approved the final version of the manuscript to be published.

Funding

This research was supported by JSPS KAKENHI Grant Numbers JP20 K21715, JP23 K09551.

Data availability

The datasets used and analyzed for the current study are not publicly available for data protection reasons.

Declarations

Ethics approval and consent to participate

The ethics committee of Kochi Medical School approved the study protocol and waived the need for patient informed consent owing to the retrospective nature of the study (Reference Number: 2021–89). The study was conducted in accordance with the principles of the Declaration of Helsinki of 1975 (as amended in 1983).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Center for Innovative and Translational Medicine, Kochi Medical School, Kochi University, Kohasu, Oko-Cho, Nankoku-Shi, Kochi 783-8505, Japan. [2]Japanese Red Cross Kochi Hospital, 1-4-63-11 Hadaminamimachi, Kochi-Shi, Kochi 780-8562, Japan. [3]Center of Medical Information Science, Kochi Medical School, Kochi University, Kohasu, Oko-Cho, Nankoku-Shi, Kochi 783-8505, Japan. [4]Department of Neuropsychiatry, Kochi Medical School, Kochi University, Kohasu,Oko-Cho, Nankoku-Shi, Kochi 783-8505, Japan.

Received: 18 September 2024 Accepted: 24 April 2025
Published online: 09 May 2025

References

1. Alzheimer's Disease International. World Alzheimer report 2022. 2022. https://www.alzint.org/resource/world-alzheimer-report-2022/. Accessed 30 Jul 2024.
2. Ohno Y, Kunisawa N, Shimizu S. Antipsychotic treatment of behavioral and psychological symptoms of dementia (BPSD): management of extrapyramidal side effects. Front Pharmacol. 2019;10:1045.
3. Bessey LJ, Walaszek A. Management of behavioral and psychological symptoms of dementia. Curr Psychiatry Rep. 2019;21:66.
4. 2021 Alzheimer's disease facts and figures. Alzheimers Dement. 2021;17:327–406.
5. Alzheimer's Disease International. World Alzheimer report 2021. 2021. https://www.alzint.org/resource/world-alzheimer-report-2021/. Accessed 30 Jul 2024.
6. Statistics Bureau, Ministry of Internal Affairs and Communications of Japan. Statistical handbook of Japan 2022. 2022. https://www.stat.go.jp/english/data/handbook/pdf/2022all.pdf. Accessed 30 Jul 2024.
7. Nakahori N, Sekine M, Yamada M, Tatsuse T, Kido H, Suzuki M. Future projections of the prevalence of dementia in Japan: results from the Toyama Dementia Survey. BMC Geriatr. 2021;21:602.
8. Zhang XX, Tian Y, Wang ZT, Ma YH, Tan L, Yu JT. The epidemiology of Alzheimer's disease modifiable risk factors and prevention. J Prev Alzheimers Dis. 2021;8:313–21.
9. Japanese Society of Neurology. Clinical practice guideline for dementia 2017. 2017. https://neurology-jp.org/guidelinem/dementia/index.html. Accessed 30 Jul 2024.
10. United States Food and Drug Administration. FDA's decision to approve new treatment for Alzheimer's disease. 2021. https://www.fda.gov/drugs/news-events-human-drugs/fdas-decision-approve-new-treatment-alzheimers-disease. Accessed 30 Jul 2024.
11. Biogen Inc. Biogen to realign resources for Alzheimer's disease franchise. 2024. https://investors.biogen.com/news-releases/news-release-details/biogen-realign-resources-alzheimers-disease-franchise. Accessed 1 Aug 2024.
12. United States Food and Drug Administration. FDA converts novel Alzheimer's disease treatment to traditional approval. 2023. https://www.fda.gov/news-events/press-announcements/fda-converts-novel-alzheimers-disease-treatment-traditional-approval. Accessed 30 Jul 2024.
13. EISAI Co., Ltd. "LEQEMBI® intravenous infusion" (Lecanemab) approved for the treatment of Alzheimer's disease in Japan. 2023. https://www.eisai.com/news/2023/news202359.html?_gl=1*1griwks*_ga*MTI1MTg1NzUwMS4xNzA3ODk1ODg0*_ga_X1FWS6YR87*MTcxMDQ4OTE1MS4zLjEuMTcxMDQ4OTE4NC4yNy4wLjA. Accessed 30 Jul 2024.
14. van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen C, Gee M, et al. Lecanemab in early Alzheimer's disease. N Engl J Med. 2023;388:9–21.
15. EISAI Co., Ltd. EISAI's approach to U.S. pricing for LEQEMBI™ (LECANEMAB), a treatment for early Alzheimer's disease, sets forth our concept of "societal value of medicine" in relation to "price of medicine". 2023. https://www.eisai.com/news/2023/news202302.html. Accessed 30 Jul 2024.
16. Jönsson L, Wimo A, Handels R, Johansson G, Boada M, Engelborghs S, et al. The affordability of lecanemab, an amyloid-targeting therapy for Alzheimer's disease: an EADC-EC viewpoint. Lancet Reg Health Eur. 2023;29:100657.
17. Deckers K, van Boxtel MPJ, Schiepers OJG, de Vugt M, Muñoz Sánchez JL, Anstey KJ, et al. Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. Int J Geriatr Psychiatry. 2015;30:234–46.
18. Beydoun MA, Beydoun HA, Gamaldo AA, Teel A, Zonderman AB, Wang Y. Epidemiologic studies of modifiable factors associated with cognition and dementia: systematic review and meta-analysis. BMC Public Health. 2014;14:643.
19. Plassman BL, Williams JW Jr, Burke JR, Holsinger T, Benjamin S. Systematic review: factors associated with risk for and possible prevention of cognitive decline in later life. Ann Intern Med. 2010;153:182–93.
20. Sofi F, Valecchi D, Bacci D, Abbate R, Gensini GF, Casini A, et al. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. J Intern Med. 2011;269:107–17.
21. Ogura H, Sagara E, Yamamoto K, Furutani H, Kitazoe Y, Takeda Y. Analysis of the online order entry process in an integrated hospital information system. Comput Biol Med. 1985;15:381–93.
22. Otomo K, Horino T, Miki T, Kataoka H, Hatakeyama Y, Matsumoto T, et al. Serum uric acid level as a risk factor for acute kidney injury in hospitalized patients: a retrospective database analysis using the integrated medical information system at Kochi Medical School hospital. Clin Exp Nephrol. 2016;20:235–43.
23. Watanabe S, Kawano T, Horino T, Matsumoto T, Nagata K, Hatakeyama Y, et al. Influence of hyperuricemia treatment on postoperative acute kidney injury among hyperuricemia patients: a single-center retrospective database analysis. BMC Res Notes. 2019;12:756.
24. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984;34:939–44.
25. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7:263–9.
26. Weber GM, Kohane IS. Extracting physician group intelligence from electronic health records to support evidence based medicine. PLoS ONE. 2013;8: e64933.
27. Ikejima C, Hisanaga A, Meguro K, Yamada T, Ouma S, Kawamuro Y, et al. Multicentre population-based dementia prevalence survey in Japan: a preliminary report. Psychogeriatrics. 2012;12:120–3.
28. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41–55.
29. Rosenbaum PR. Model-based direct adjustment. J Am Stat Assoc. 1987;82:387–94.
30. Rubin DB. The use of propensity scores in applied Bayesian inference. In: Bernardo JM, DeGroot MH, Lindley DV, Smith AFM, editors. Bayesian statistics 2. North-Holland: Elsevier Science Publisher BV; 1985. p. 463–72.
31. R4.1.3 (R Core Team (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021. URL https://www.R-project.org/.
32. Greifer N. WeightIt: weighting for covariate balance in observational studies. R package version 0.13.1; 2022. https://CRAN.R-project.org/package=WeightIt.
33. Greifer N. Cobalt: covariate balance tables and plots. R package version 4.3.2; 2022. https://CRAN.R-project.org/package=cobalt.
34. World Health Organization. Serum and red blood cell folate concentrations for assessing folate status in populations. 2012. https://iris.who.int/bitstream/handle/10665/75584/WHO_NMH_NHD_EPG_12.1_eng.pdf?sequence=1. Accessed 19 Feb 2025.
35. Nakamura K. Aniracetam: its novel therapeutic potential in cerebral dysfunctional disorders based on recent pharmacological discoveries. CNS Drug Rev. 2002;8:70–89.
36. Ohtsuki K, Abe Y, Shimoyama Y, Furuya T, Munakata H, Takasaki C. Separation of phospholipase A2 in Habu snake venom by glycyrrhizin (GL)-affinity column chromatography and identification of a GL-sensitive enzyme. Biol Pharm Bull. 1998;21:574–8.
37. Dhapola R, Hota SS, Sarma P, Bhattacharyya A, Medhi B, Reddy DH. Recent advances in molecular pathways and therapeutic implications targeting neuroinflammation for Alzheimer's disease. Inflammopharmacology. 2021;29:1669–81.
38. Wang C, Zong S, Cui X, Wang X, Wu S, Wang L, et al. The effects of microglia-associated neuroinflammation on Alzheimer's disease. Front Immunol. 2023;14:1117172.
39. Gibson GE, Hirsch JA, Fonzetti P, Jordan BD, Cirio RT, Elder J. Vitamin B1 (thiamine) and dementia. Ann N Y Acad Sci. 2016;1367:21–30.
40. Perrin RJ, Fagan AM, Holtzman DM. Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. Nature. 2009;461:916–22.
41. Furst AJ, Rabinovici GD, Rostomian AH, Steed T, Alkalay A, Racine C, et al. Cognition, glucose metabolism and amyloid burden in Alzheimer's disease. Neurobiol Aging. 2012;33:215–25.
42. Ministry of Internal Affairs and Communications. Population estimates (in Japanese, As of October 1, 2022). https://www.stat.go.jp/data/jinsui/2022np/pdf/2022np.pdf. Accessed 7 Feb 2025.
43. Kochi Prefecture. Transportation Access (in Japanese). https://www.pref.kochi.lg.jp/richi/about/access.html. Accessed 15 Feb 2025.
44. American Psychiatric Association. Diagnostic & Statistical Manual of Mental Disorders: DSM-5. Amer Psychiatric Pub Inc; 2013.

45. Cannon JA, Moffitt P, Perez-Moreno AC, Walters MR, Broomfield NM, McMurray JJV, et al. Cognitive Impairment and Heart Failure: Systematic Review and Meta-Analysis. J Card Fail. 2017;23:464–75.
46. Zlokovic BV. Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders. Nat Rev Neurosci. 2011;12:723–38.
47. Kazui H, Yoshiyama K, Kanemoto H, Suzuki Y, Sato S, Hashimoto M, et al. Differences of Behavioral and Psychological Symptoms of Dementia in Disease Severity in Four Major Dementias. PLoS ONE. 2016;11: e0161092.
48. Hoffmann J, Busse S, von Hoff F, Borucki K, Frodl T, Busse M. Association Between Homocysteine and Vitamin Levels in Demented Patients. J Alzheimers Dis. 2021;81:1781–92.
49. Malouf R, Areosa Sastre A. Vitamin B12 for cognition. Cochrane Database Syst Rev. 2003;3:CD004394.
50. Robinson N, Grabowski P, Rehman I. Alzheimer's disease pathogenesis: Is there a role for folate? Mech Ageing Dev. 2018;174:86–94.
51. Shen L, Ji HF. Associations between Homocysteine, Folic Acid, Vitamin B12 and Alzheimer's Disease: Insights from Meta-Analyses. J Alzheimer's Dis. 2015;46:777–90.
52. Ravaglia G, Forti P, Maioli F, Martelli M, Servadei L, Brunetti N, et al. Homocysteine and folate as risk factors for dementia and Alzheimer disease. Am J Clin Nutr. 2005;82:636–43.
53. Quinlan P, Horvath A, Wallin A, Svensson J. Low serum concentration of free triiodothyronine (FT3) is associated with increased risk of Alzheimer's disease. Psychoneuroendocrinology. 2019;99:112–9.
54. Dolatshahi M, Salehipour A, Saghazadeh A, Sanjeari Moghaddam H, Aghamollaii V, Fotouhi A, et al. Thyroid hormone levels in Alzheimer disease: a systematic review and meta-analysis. Endocrine. 2023;79:252–72.
55. Begic E, Hadzidedic S, Kulaglic A, Ramic-Brkic B, Begic Z, Causevic M. SOMAscan-based proteomic measurements of plasma brain natriuretic peptide are decreased in mild cognitive impairment and in Alzheimer's dementia patients. PLoS ONE. 2019;14: e0212261.
56. Hu WT, Holtzman DM, Fagan AM, Shaw LM, Perrin R, Arnold SE, et al. Plasma multianalyte profiling in mild cognitive impairment and Alzheimer disease. Neurology. 2012;79:897–905.
57. Kondziella D, Göthlin M, Fu M, Zetterberg H, Wallin A. B-type natriuretic peptide plasma levels are elevated in subcortical vascular dementia. NeuroReport. 2009;20:825–7.
58. Marksteiner J, Imarhiagbe D, Defrancesco M, Deisenhammer EA, Kemmler G, Humpel C. Analysis of 27 vascular-related proteins reveals that NT-proBNP is a potential biomarker for Alzheimer's disease and mild cognitive impairment: a pilot-study. Exp Gerontol. 2014;50:114–21.
59. Nagata T, Ohara T, Hata J, Sakata S, Furuta Y, Yoshida D, et al. NT-proBNP and Risk of Dementia in a General Japanese Elderly Population: The Hisayama Study. J Am Heart Assoc. 2019;8: e011652.
60. Paudel YN, Angelopoulou E, Semple B, Piperi C, Othman I, Shaikh MF. Potential neuroprotective effect of the HMGB1 inhibitor glycyrrhizin in neurological disorders. ACS Chem Neurosci. 2020;11:485–500.
61. Ministry of Health, Labor and Welfare. List of Drug Price Listings and Information on Generic Drugs: Injectable Drugs (in Japanese, effective February 1, 2025). 2025. https://www.mhlw.go.jp/topics/2024/04/dl/tp20250201-01_02.pdf. Accessed 7 Feb 2025.
62. Sheng B, Wang X, Su B, Lee HG, Casadesus G, Perry G, et al. Impaired mitochondrial biogenesis contributes to mitochondrial dysfunction in Alzheimer's disease. J Neurochem. 2012;120:419–29.
63. Wang X, Su B, Siedlak SL, Moreira PI, Fujioka H, Wang Y, et al. Amyloid-beta overproduction causes abnormal mitochondrial dynamics via differential modulation of mitochondrial fission/fusion proteins. Proc Natl Acad Sci U S A. 2008;105:19318–23.
64. Wang X, Su B, Lee HG, Li X, Perry G, Smith MA, et al. Impaired balance of mitochondrial fission and fusion in Alzheimer's disease. J Neurosci. 2009;29:9090–103.
65. Zhu X, Chen C, Ye D, Guan D, Ye L, Jin J, et al. Diammonium glycyrrhizinate upregulates PGC-1α and protects against Aβ1-42-induced neurotoxicity. PLoS ONE. 2012;7: e35823.
66. Tuppo EE, Arias HR. The role of inflammation in Alzheimer's disease. Int J Biochem Cell Biol. 2005;37:289–305.
67. Zhao H, Wang SL, Qian L, Jin JL, Li H, Xu Y, et al. Diammonium glycyrrhizinate attenuates $A\beta_{1-42}$-induced neuroinflammation and regulates MAPK and NF-κB pathways in vitro and in vivo. CNS Neurosci Ther. 2013;19:117–24.
68. Fujita K, Motoki K, Tagawa K, Chen X, Hama H, Nakajima K, et al. HMGB1, a pathogenic molecule that induces neurite degeneration via TLR4-MARCKS, is a potential therapeutic target for Alzheimer's disease. Sci Rep. 2016;6:31895.
69. Kong ZH, Chen X, Hua HP, Liang L, Liu LJ. The oral pretreatment of glycyrrhizin prevents surgery-induced cognitive impairment in aged mice by reducing neuroinflammation and Alzheimer's-related pathology via HMGB1 inhibition. J Mol Neurosci. 2017;63:385–95.
70. Tabuchi M, Imamura S, Kawakami Z, Ikarashi Y, Kase Y. The blood-brain barrier permeability of 18β-glycyrrhetinic acid, a major metabolite of glycyrrhizin in Glycyrrhiza root, a constituent of the traditional Japanese medicine yokukansan. Cell Mol Neurobiol. 2012;32:1139–46.
71. Jourdan JP, Bureau R, Rochais C, Dallemagne P. Drug repositioning: a brief overview. J Pharm Pharmacol. 2020;72:1145–51.
72. Research project for the establishment of an NDB research system for health policy and other purposes through 6 NC collaboration. About the NDB. https://ndb6nc.ncgm.go.jp/eng/outline/index.html. Accessed 30 Jul 2024.
73. Ministry of Health, Labor and Welfare. A manual for those considering using NDB (in Japanese). 2025. https://www.mhlw.go.jp/content/12400000/001427722.pdf. Accessed 4 Mar 2025.
74. Ge Q, Chen L, Tang M, Zhang S, Liu L, Gao L, et al. Analysis of mulberry leaf components in the treatment of diabetes using network pharmacology. Eur J Pharmacol. 2018;833:50–62.
75. Pang XC, Kang D, Fang JS, Zhao Y, Xu LJ, Lian WW, et al. Network pharmacology-based analysis of Chinese herbal Naodesheng formula for application to Alzheimer's disease. Chin J Nat Med. 2018;16:53–62.
76. Zhang M, Zheng H, He J, Zhang M. Network pharmacology and in vivo studies reveal the neuroprotective effects of paeoniflorin on Alzheimer's disease. Heliyon. 2023;9: e21800.

Publisher's Note