



RIETI Discussion Paper Series 21-J-003

アロマセラピーは、健常高齢者の認知機能改善に効果があるか？ ーランダム化比較試験による検証ー

宗 未来
東京歯科大学

小杉 良子
慶應義塾大学

新生 暁子
順天堂大学

小田木 友依
慶應義塾大学

腰 みさき
慶應義塾大学

橋本 空
ユナイテッド・ヘルスコミュニケーション株式会社

関沢 洋一
経済産業研究所

斎藤 文恵
慶應義塾大学

小西 海香
慶應義塾大学

森 恵莉
慈恵医科大学

船山 道隆
足利赤十字病院

田淵 肇
慶應義塾大学

三村 将
慶應義塾大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

アロマセラピーは、健常高齢者の認知機能改善に効果があるか？¹

ーランダム化比較試験による検証ー

宗 未来^{#1#2}、小杉 良子^{#2}、新生 暁子^{#3}、小田木 友依^{#2}、腰 みさき^{#2}、橋本 空^{#4}、関沢 洋一^{#5}、
斎藤 文恵^{#2}、小西 海香^{#2}、森 恵莉^{#6}、船山 道隆^{#7}、田淵 肇^{#2}、三村 将^{#2}

^{#1} 東京歯科大学 ^{#2} 慶應義塾大学 ^{#3} 順天堂大学 ^{#4} ユナイテッド・ヘルスコミュニケーション株式会社 ^{#5} 経済産業研究所 ^{#6} 慈恵医科大学 ^{#7} 足利赤十字病院

要 旨

背景：超高齢社会を迎え、エビデンスに裏付けされた認知機能改善の開発が急務である。アルツハイマー型を含む多くの認知症では、記憶より先に嗅覚が衰えるが、嗅覚刺激は逆に記憶を司る海馬の神経新生を促すため、認知機能の改善が期待される。しかし、厳密な臨床試験での証明は事実上皆無である。本研究は、アロマセラピーで高齢者の認知機能が改善するかをランダム化比較試験で検証した。

研究方法：健常高齢者に、朝晩2時間ずつアロマシールを洋服に貼るアロマ群60名と、アロマの代わりにエタノールを使うプラセボ群59名に分け、12週間の介入を行った。主要評価項目は注意機能検査であるPASATで、副次的評価項目として記憶等の他の認知機能検査や対人信頼度等の心理社会的評価も行った。

結果：PASAT-2秒条件で、アロマ群はプラセボ群に対して有意な改善を認めた($p=0.0223$, Hedges' $g=0.44$)。他の殆どの認知機能や心理学評価では差を認めなかったが、ウェルビーイングと嗅覚(同定能力)ではプラセボ群が優っていた。

結論：脳機能の中でも加齢に伴いまず先に低下する注意は、記憶や他の認知機能よりも認知症や軽度認知障害(MCI)を有する高齢者の自立生活破綻や危険運転等への影響が大きいことから、近年重視されており、アロマセラピーにその注意改善効果が示されたことの臨床的意義は大きい。一方で他の大半の認知機能や心理社会的変数への効果は示されず、対象が健常高齢者であるため天井効果により伸びしろが少なかったことも一因と示唆された。今後の検証が望まれる。

キーワード：アロマセラピー、認知症、認知機能、嗅覚刺激、ランダム化比較試験(RCT)

JEL classification: I10

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び(独)経済産業研究所としての見解を示すものではありません。

¹本稿は、独立行政法人経済産業研究所(RIETI)におけるプロジェクト「エビデンスに基づく医療に立脚した医療費適正化策や健康経営のあり方の探求」の成果の一部である。

1. 背景

認知症は、高齢者本人に多大な負荷を生じさせるだけでなく、医療費、介護費、社会保障費、家族による無償の介護等の形で、家族のみならず社会全体にまで広くその負担が及ぶ。日本における認知症の社会的コストは、2014年の時点で14.5兆円に上り、2060年には24.3兆円と現在の約1.6倍にまで膨張すること、生産労働人口一人あたりの社会的コストに至っては、現在の2.8倍以上になることが報告されており[1]、高齢者の認知症問題が今後一層、本人、家族、そして社会に甚大な影響を及ぼすことは明らかである。しかしながら、現在まで、アルツハイマー型認知症の治療薬に関する臨床試験には時間がかかっており、薬物療法の限界が示唆されている[2]。現状では、この問題の現実的な解決策として、「疾病予防による健康寿命の延伸化」以外に、「リハビリテーションによる機能回復」が期待されている。

「認知症は、記憶より前に嗅覚が損なわれる」との報告が近年、認められる。2017年春にはエーザイ株式会社が米ペンシルバニア大学の開発した技術を基に、「認知症の兆候発見嗅覚検査キット」を販売するなど、認知症と嗅覚障害については注目が集まっている[3]。

学術的には、もっとも多い認知症タイプであるアルツハイマー型認知症では、記憶に大きく関係する海馬領域よりも先に、より外側にあり嗅覚を主に司る「嗅内皮質」が冒される（神経原線維変化が生じる）ことが多い[4, 5]。そのために、記憶障害よりも先に嗅覚低下がアルツハイマー型認知症では症状として出現する可能性がある。最新のメタ解析（2019）でも、アルツハイマー型認知症者ではMCI（軽度認知障害）よりも嗅覚同定能力の障害が大きいと報告されている[4]。嗅覚障害を有する高齢者は、アルツハイマー型認知症になるリスクが高いことも報告されており[6, 7]。現在、その感度・特異度とも約80%であることも示されている。また、アルツハイマー型認知症と並んで四大認知症のひとつであるレビー小体型認知症（DLB：Dementia with Lewy bodies）でも、早期から嗅覚障害が認められることが知られている[8]。嗅覚障害は、DLBでは全経過中90%以上に認められ、類縁疾患ともいえるパーキンソン病も含めたレビー小体病では運動症状が出現する以前から嗅覚認知に関連する嗅球、嗅索や扁桃核にレビー小体病理変化が出現することが認められている。

また、ヒトにおける嗅神経と海馬ではその生涯にわたり神経新生という神経細胞の再生能力があることが分かっており、特に再生能力が高い嗅神経を効果的に刺激し続けることで、嗅神経細胞の神経新生が活性化され、更にその刺激が海馬にも波及し、海馬や周囲の神経細胞の神経新生が促進されることが示されている。その結果、効果的な嗅覚刺激が、認知機能改善に寄与するという仮説が提唱されている[9-11]。

しかし、我々の知る限り、効果が証明されている認知症患者への嗅覚刺激研究の多くは焦燥や興奮といった行動・心理面における周辺症状（BPSD）減少の報告が中心である。国際的に信頼性の高いコクラン・レビュー（2014）[12]によっても、解析対象となる研究自体が不十分としながらも認知症者の焦燥や問題行動への有意な効果がメタ解析で示されている。一方で、記憶力などの認知症で問題となる中核的な認知機能の改善の有無については殆ど検証に耐える研究がなされておらず、結論が出ていない。上記のコクラン・レビュー（2014）[12]でも、認知症患者の認知機能に対して系統的レ

ビューの解析対象とはされたものの、文献検索の結果、報告が皆無のために解析対象外の扱いとなっている。

一方、国内のウェブサイトなどを見ても、十分な医科学的根拠に欠けたまま介護福祉サービスの現場などではアロマセラピーの認知症予防もしくは改善効果が喧伝され、ブームと言えるほど広く一般に普及している。そのような状況の中で、根拠に基づく科学的介入を目指すべき立場からも十分な検証が望まれる。

これまでの先行研究では、中核症状である認知機能自体の嗅覚刺激による改善についての検討は Jimbo ら[13]による報告がよく知られている。アルツハイマー型認知症 10 名を含む軽-中等度の認知症または認知症予備軍 (MCI) の 28 名に対して 28 日間の午前 (ローズマリーとレモン) と午後 (ラベンダーとオレンジ) の芳香浴による前後比較で、自己に関する見当識 (自分の姓名・職業・年齢・誕生日の正しい認識) の有意な改善などが示されている。これらの結果を受ける形で、2020 年にアルツハイマー型認知症やその予備軍である MCI を有する平均年齢 80 歳の高齢者に対して、アロマバスでの認知機能への効果検証が、ランダム化比較試験によるはじめての試みとして報告された[14]。しかし、結果は認知機能、睡眠状態、嗅覚機能のすべてにおいて変化は認められないというものであった。

認知症患者のアロマセラピーでは、特に懸念されている点として認知症患者では高度に嗅覚が障害されていることが多く、認知機能改善の有無の議論の前にそもそもその前提であるアロマセラピーでの介入がどこまで嗅覚刺激として働きかけられるのかという点が挙げられる。そのため、認知症患者に対するアロマセラピーの効果は限界が多いという見解さえ示されている[15]。実際に、Kouzuki らの研究でも[14]、参加者全員の嗅覚機能は、簡易嗅覚同定テストである“においスティックテスト (OSIT-J)”において、全 12 問のうち平均正答数 2-5 個と著しく低いものであった (健常成人は 9/12 以上の正答数で正常と評価される)。そのため、認知症患者へのアロマセラピーの認知機能への効果検証には、①前提となる嗅覚刺激が本当になされているのかという点と、②刺激された嗅覚がそれによって実際に認知機能を改善させるのかという点という、2 段階の検証が明確化されなければならない。

たとえば、認知症のない 65 歳以上の健常高齢者においてさえも、においスティックでの同定能力テストで既に平均で 9 点程度に落ちていると報告される。Moss らは[16, 17]、嗅覚機能が低下した高齢者での検討ではなく、そもそも脳神経細胞の可塑性が高く、かつ嗅覚が正常な平均年齢 20 代半ばの若者 44 名に対して、アロマセラピーで認知機能がどの程度改善するかについてランダム化比較試験で検証している。その結果、興奮系のローズマリーやペパーミントは記憶力の改善に寄与する一方で、鎮静系のラベンダーやイランイランは記憶力を低下させ、さらに 4 つの精油のすべてが処理速度を低下させることを報告している。国内では、川畑と上出が[18]やはり平均年齢で 20 代前半の 45 名の若者に対して、ラベンダー群、ブラックペッパー群、プラセボ群の 3 群でのランダム化比較試験を行い、注意機能を主として評価する PASAT (Paced Auditory Serial Addition Task) では有意差は認めなかったものの、遂行能力を評価するロンドン塔課題によってラベンダー群だけがプラセボ群に有意な改善を認めたと報告

している。しかし、上記研究はすべて平均年齢が 20 代を対象としており、高齢者にはそのまま適応できない。

共同研究者である田渕らは、健常高齢者に対して比較対照を置かないシングルアームでの予備的介入研究において朝夕のアロマセラピー（専用ペンダントに精油を入れ香りを吸入する）を 8 週間にわたって行い（2015）、その結果、①記憶（言語性）の評価検査である RAVLT（Rey Auditory Verbal Learning Test）においては即時再生総数（ $p<.001$ ）、干渉後課題（ $p=.008$ ）、遅延再生（ $p<.001$ ）、②注意機能の評価検査である PASAT においては 2 秒課題（ $p<.001$ ）、1 秒課題（ $p<.001$ ）の有意な改善効果を認めた[19]。しかし、これらの結果は、どちらも前向きシングルアーム研究であり、偽薬効果や学習効果によるバイアスの可能性が否定できない。また先行研究でもランダム化比較試験による健常高齢者に対する認知機能の改善に関する報告は我々の知る限り国際的にも認められないことから、ランダム化比較試験による厳密な効果検証が求められている。

また、仮に健常高齢者においても効果が認められたとして、将来的に顕著な嗅覚低下が想定される高齢者へのアロマセラピーの普及という観点からは、嗅覚機能低下への対応も不可欠である。近年、嗅覚機能の低下した者であっても、精油を嗅ぐリハビリテーションによって嗅覚は改善されるというエビデンスが示されている[20]。さらに、Jimbo らの報告では精油の濃度を高めることでより高い認知機能改善効果が得られる可能性が示唆されており[13]、一方でランダム化比較試験では十分な効果がこれまで得られてこなかったことも踏まえて、アロマセラピーとしては決して極端なものではないか、時間・頻度ともに先行研究と比較し、従来の認知機能への働きかけや嗅覚訓練での方法論より相対的に高強度となる精油を用いた介入の効果検証が有用であろうとの仮説を我々は有した[14]。

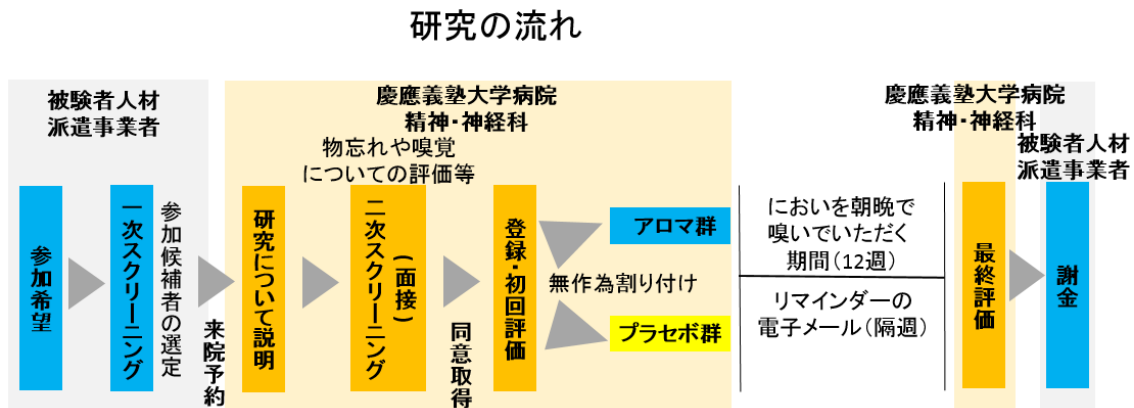
加えて、我が国における健常高齢者に対して精油を嗅がせるリハビリテーションが、認知機能や嗅覚機能だけでなく、従来報告されている心理社会的状態にもどのように影響を与えているのかといった点についても、質の高いランダム化比較試験で検証されることが望まれている。そのため、本研究ではそれらについて実際にランダム化比較試験によって検証を行った。

2. 方法（図 1）

(1)倫理審査

本研究は慶應義塾大学医学部倫理委員会の承認を受けて行われ、UMIN 臨床試験登録システムに登録された（ID= UMIN000032936）。書面同意が全参加者から取得された。

図1 依頼方法・研究全体の流れフローチャート



(2)参加者

本研究の参加者は、2018年の8月～12月末にかけて被験者人材派遣業務を行う3Hメディソリューション株式会社（以下、「派遣業者」と呼ぶ）のモニターとして登録している人々に募集をかけることによって行われた。筆者らが所属する慶應義塾大学医学部が派遣業者と委託契約を締結することによって、以下の手順に従って本研究は遂行された。

まず、派遣業者における登録者に対して参加募集が行われ、参加希望者の中で、本人からの自己申告によって物忘れで大きな問題を有しておらず、選択基準および除外基準を満たすことを事業者によって確認された者が、一次スクリーニングとして本研究の対象候補者として選定された（記憶力の低い者は、事業者独自に設定された質問式の簡易スクリーニング検査に基づき除外された）。情報バイアスを回避するために研究終了後まで、「アロマ」といった言葉は使用されず、あくまでも「におい刺激による認知機能の改善効果検証」とだけ被験者には伝えられた。

一次スクリーニングの条件を満たした者は、派遣業者を通じて慶應義塾大学病院にアポイントがとられた。指定の来訪日時に研究者によって説明文書に基づき、本研究参加および研究内容などについて口頭および文書にて十分に説明され、本研究の内容を十分に理解した志願者に対し、研究者により選択基準および除外基準を満たすことの確認〔MMSEでの神経心理学検査を実施し認知機能に問題が無いこと（MMSEが24点以上）の確認や簡易嗅覚検査を含む〕が二次スクリーニングとしてなされた。

その後、二次スクリーニングで選択条件を満たした者に対して、本研究参加について、本人の自由意志による同意を書面にて得た。選択基準は以下の通りであった。

包有基準：

- ・ 65歳以上 80歳未満（同意取得時年齢）の者
- ・ 同意取得後、指定された期日に必要な検査が実施可能と判断できる者
- ・ 同意取得に先立って実施された認知機能テスト（MMSE）において問題がない者（24点以上）

- ・同意取得に先立って実施された簡易嗅覚同定検査（においスティック:OSIT-J）において正常な者
- ・本研究参加の本人同意が文書で得られる者

除外基準：

- ・喫煙以外の物質使用障害の既往または治療中の者
- ・てんかんの既往または治療中の者
- ・教育歴が 6 年以下の者
- ・抗うつ剤、抗精神薬、長期にわたる睡眠薬を含む催眠鎮静剤（抗不安薬を含む）の投与を受けている者
- ・認知機能に影響する精神・神経疾患の既往（統合失調症、気分障害、不安障害など）を有する者 [M.I.N.I. (Mini-International Neuropsychiatric Interview) で除外]
- ・重篤な合併症（肝障害、腎障害、内分泌疾患など）を有する者
- ・嗅覚障害が生じる疾患を有する者
- ・日常的にアロマセラピーを嗜んでいる者
- ・日常的に匂いの強い柔軟剤や芳香剤、香水を使用している者

上記選択条件を満たした参加者は、同意取得後すみやかに、属性など基本情報の聴取および身長・体重の測定後、心理検査・神経心理学的検査と精神症状評価検査が実施された。

(3)標本数と無作為化

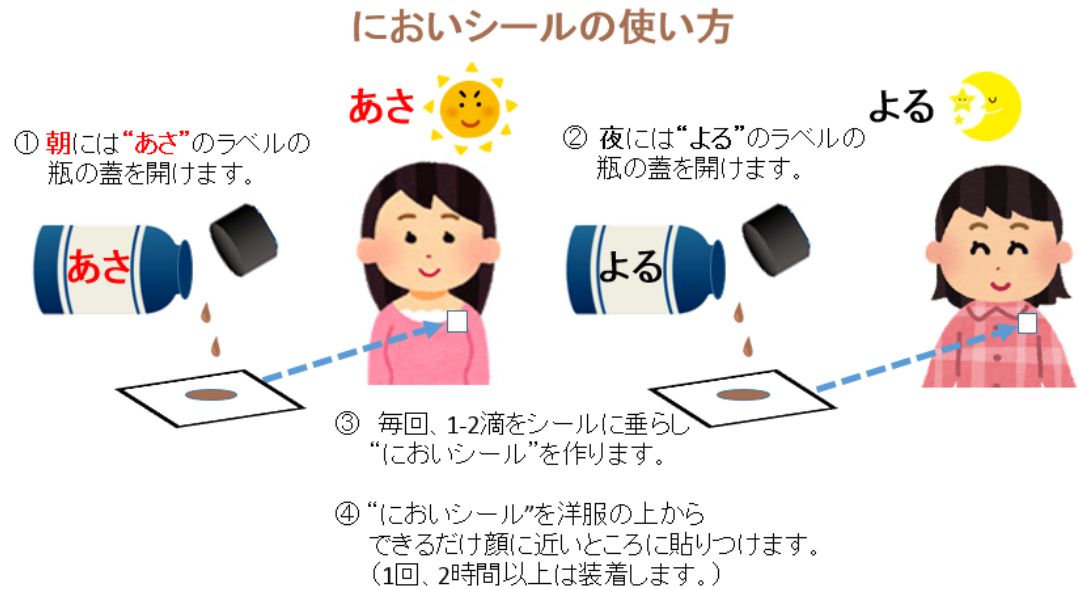
標本数は各群 60 例の計 120 例とした。算出の根拠は効果量 0.15、 α を 0.05、検出力を 0.9 として計 100 名が必要であり、そこに推定脱落率を約 20%として、計 120 名（各群 60 名）と算出された。算出ソフトは、GPower 3.1 を使用した。

ランダム化比較試験の手順は、CONSORT 2010 チェックリストに従って実施された。無作為化の方法は、本研究とは独立の第三者による中央割り付け方式で行われた。適格性のある研究協力者を介入群と対照群とに同じ割合で単純割り付けとした。割り付け者は、リサーチ・アシスタントにこのリストを送り、リサーチ・アシスタントは安全な場所でこのリストを管理し、割り付け隠蔽化が可能なスプレッド・シートを作成した。リサーチ・アシスタントは適格性を有する参加者のリストを作成し、各参加者がどの群に該当するかを確認し、上記スプレッド・シートに入力し、このスプレッド・シートに従って番号が紐付けされた外見上は識別不能のにおい液の瓶が研究実施者より研究協力者に手渡された。

(4)介入

無作為割付によって、精油による嗅覚曝露が行われるアロマ群と、精油の代わりに無水エタノールを使ったプラセボ群の二群に分けられた。両群には、それぞれどちらの液かは伝えられず、割付けられた群のにおい液が入った同じ外観のボトルが朝夕用に2本ずつ日数分手渡された。それぞれを12週にわたり1-2滴スポイトに毎回各自でにおい液を浸したパッチ（対照群はエタノール）をシールで洋服に接着し、1日2回朝晩1回2時間以上の嗅覚刺激を行うように指示がなされた。また毎回、実施するたびにアドヒアランスシートにその有無を入力することも課された（図2）。

図2 嗅覚曝露のイメージ



12（±2）週間のエクササイズ期間が終了後に、第2回目の評価指標のアセスメントが行われた。再度心理検査（神経心理検査・精神症状評価検査）および簡易嗅覚検査の測定が実施され評価された。介入期間は2019年1月から6月までであった（表1）。

表1 試験手続きの流れ

ID	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	wk21	wk22
101～112	◎	●	●	●	●	●	●	●	●	●	●	●	○									
213～224		◎	●	●	●	●	●	●	●	●	●	●	●	○								
325～348			◎	●	●	●	●	●	●	●	●	●	●	○								
449～460				◎	●	●	●	●	●	●	●	●	●	●	○							
561～572					◎	●	●	●	●	●	●	●	●	●	●	○						
673～684						◎	●	●	●	●	●	●	●	●	●	●	○					
785～796							◎	●	●	●	●	●	●	●	●	●	●	○				
予備								◎	●	●	●	●	●	●	●	●	●	●	○			
予備									◎	●	●	●	●	●	●	●	●	●	●	○		
予備										◎	●	●	●	●	●	●	●	●	●	●	○	

（注）◎は事前説明、同意取得、介入前検査。●は介入期間。○は介入後検査。

アロマ群：

本研究で使用されたのは市販のブレンド精油で、朝用と夕用をそれぞれ別々の種類の精油で構成された2種のブレンド精油が使用された。朝用には、交感神経系を活性化する目的とし、ローズマリーシネオール (*Rosmarinus officinalis ct.cineole*)、ペパーミント (*Mentha piperita*)、ティートリー (*Melaleuca alternifolia*)、サイプレス (*Cupressus sempervirens*) の4種類の精油をブレンドしたにおい液を用いた。夜用としては、副交感神経系を刺激し鎮静作用を目的とし、ラバンジン (*Lavandula hybrida*)、スイートオレンジ (*Citrus sinensis*)、プチグレン (ビターオレンジ葉) (*Citrus aurantium*)、ラベンダー・ファイン (*Lavandula angustifolia*)、スイートマジョラム (*Origanum majorana*)、レモングラス (*Cymbopogon citratus*)、バジル (*Ocimum basilicum*) の7種類の精油をブレンドしたにおい液を用いた。それぞれのブレンド精油はすべて原液 (濃度は100%希釈なし) であった。

プラセボ群：

プラセボ群では、無水エタノールがにおい液として使用された。

(5)研究参加者への謝礼

研究参加者への謝礼として、初回指定日のみに来訪して規定の研究協力事項の完遂が確認された研究協力者に2,000円、初回と2回目の両指定日に来訪して規定の研究協力事項の完遂が確認された研究協力者には10,000円が派遣業者を介して提供された。

(6)評価指標

① 主要評価項目

本研究における主要評価項目は、試験開始前と試験開始12週後の注意機能の変化量の2群間比較とし、注意機能はPASAT (Paced Auditory Serial Addition Test) によって計測した。試験開始時(0週)、試験終了時(12週)の2回評価とした。なお、介入期間の評価は、予定評価日の±2週間の範囲を許容とした。

② 副次的評価項目

言語性記憶検査 (RAVLT : Rey Auditory Verbal Learning Test)、注意機能や処理速度 (SDMT : Symbol Digit Modalities Test)、全般的認知機能 (MMSE : Mini-Mental State Examination)、抑うつ (PHQ-9 : Patient Health Questionnaire-9)、不安 (GAD-7 : Generalized Anxiety Disorder-7)、機能障害 (SDS : Sheehan Disability Scale)、ウェルビーイング (CES-D : The Center for Epidemiologic Studies Depression Scale の該当4項目)、一般的信頼尺度、睡眠 (PSQI : Pittsburgh Sleep Quality Index)、簡易嗅覚同定機能検査 (においスティック : OSIT-J) について、試験開始時(0週間)、試験終了時(12週)の2

回で評価された。介入期間の評価は、予定評価日の±2週間の範囲を許容とした。また、試験終了時に、試験前に予め手渡しておいたアドヒアランスチェックシート（○X式）が回収され、香りについての好き嫌い、有益有害事象の発生についても聴取された。

③それぞれの評価尺度の概要

◆PASAT : Paced Auditory Serial Addition Task

高次の注意機能を反映するテストであり、聴覚的注意(容量や選択性、配分能力)の評価テストとして優れているとされる。検査では、ランダムに読み上げられる1-9の数字を1回ずつ足し算していく。この足し算が行われている間に次の数字が示されるため、計算をしながらも、すぐ前に示された数字を覚え、かつ次の数字を聴き取らなければならない。数字が2秒間隔で提示される2秒条件と、1秒間隔で提示される1秒条件とがあり、後者はより負荷が大きく難易度が高い（図3）。

図3 PASAT の概要

**Paced Auditory Serial Addition Test
(PASAT)**

聴覚的な分配性注意などを評価する。

※ () 内は正答

【2秒条件：練習】トラック9 7 2 3 4 2 5 8
 (9) (5) (7) (6) (7) (13)



1桁の足し算をします。自動音声で「2・8」と言ったら、足して「10」と答えてください。
そして8を覚えていて、次に6と言ったら8と6を足して14と答えてください。
このようにひとつ前の数字と新しく言う数字を足していく問題です。

◆RAVLT : Rey Auditory Verbal Learning Test

言語性記憶の評価法で、15単語を繰り返し記憶していく過程での成績を調べる。1) リストAの15単語の即時再生、2) リストAの記憶学習を5回繰り返す過程での学習効果、さらに3) リストBの単語を記憶する干渉後に、リストAで覚えているものをできるだけ想起する遅延再生（干渉後）と、4) 提示された単語がリストAにあったか否かについて答えて正答数を評価する再認で構成されている。アルツハイマー型認知症やその予備軍とされるMCIにおける記憶障害では、①学習効果〔第5試行での想起達成

数から第1試行の達成数から第1試行の達成数を引いた言語学習能力（verbal learning ability：VLA）や“5回目想起数”]と②遅延再生が特に重要であると考えられている[21-23]。実際、軽度認知障害（MCI）から将来、アルツハイマー型認知症へ進展することを予測する検査としては、他の神経心理学的検査、生物学的指標、海馬容積などの画像検査と比較しても、単独でのベストはRAVLTであったと報告されるなど優れた記憶検査として知られる[24]。

◆SDMT：Symbol Digit Modalities Test

「記号と数字の対応」に従って、提示された図形に対応する数字を90秒間でできるだけ早く正確に記入していく試験で、視覚的な分配性注意と処理速度を反映するとされる（図4）。

図4 SDMTの概要

Symbol Digit Modalities Test (SDMT)

視覚的な分配性注意などを評価する。

記号と数字の対応

(	⊖	⊢	⌈	⊥	>	+	)	÷
1	2	3	4	5	6	7	8	9

(	⊥	⊖	(	⊢	>	⊖	⌈	(	>	⊖	(	>	(	⊖
1	5	8	1	3	..	..								

ここに9つの記号があります。この記号には1、この記号は2、というように、それぞれに対応する数字が決まっています。下を見てください。記号だけ書いてあって、数字のところは空欄になっています。ここに当てはまる数字を上から探して書き入れていく問題です。



◆PHQ-9：Patient Health Questionnaire -9

PHQ-9は、大うつ病性障害等の診断のために開発された質問票で、多忙なプライマリケア医が短時間で精神疾患を診断・評価するための自記式尺度である。英国国立医療技術評価機構（NICE：National Institute of Health and Clinical Excellence）のうつ病治療のガイドラインや米国精神医学会（APA：American Psychiatric Association）は、うつ病の評価尺度としてPHQ-9を推奨している。DSM-5 (The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition)に対応する9個の質問からPHQ-9は構成されている。

◆GAD-7：Generalized Anxiety Disorder-7

GAD-7は、全般性不安障害を簡易に評価するための自記式質問票として開発されたものである。7つの質問から構成されており、過去2週間についての不安症状につい

て評価される。英国における国立医療技術評価機構（NICE：National Institute of Health and Clinical Excellence）ガイドラインでは全般的な不安評価尺度として、GAD-7を推奨している。

◆SDS：Sheehan Disability Scale

SDSは、機能障害評価尺度の一つである。この指標は疼痛評価における視覚的評価スケール（Visual Analog Scale：VAS）と同様に社会機能障害を数字で評価することができる。その簡便さと治療効果に対する正確性ゆえに、米国では特にこの指標が広まりつつある。SDS日本語版とGlobal assessment of functioning（GAF）との相関でも高い併存妥当性が示されており、SDS日本語版は高い信頼性・妥当性を有し、機能障害を測定する簡易尺度として有用であることが示されている。

◆CES-D：The Center for Epidemiologic Studies Depression Scale（ウェルビーイング指標の4項目）

CES-Dは、アメリカの国立精神保健研究所が開発したうつ病の自己評価尺度で、20問の質問から構成されている。気分や身体の状態が過去1週間にどの程度持続したかについて選択する。CES-Dは、ネガティブな項目に関する質問が16問ある一方で、「人並みのことはできると感じた」「さきゆき明るいと感じた」「うれしいと感じた」「楽しいと感じた」というポジティブな質問が4問あり、このポジティブな質問は幸福度とも称される概念である“ウェルビーイング”の指標として使われており、合計得点が高いほどウェルビーイングの程度は高いと評価される。

◆ピッツバーグ睡眠質問票（PSQI：Pittsburgh Sleep Quality Index）

PSQIは、睡眠障害の評価として広く使用されており、睡眠の質、入眠時間、睡眠時間、睡眠効率、睡眠困難、睡眠薬の使用、日中覚醒困難の7要素と質問票総合得点として算出される。信頼性、妥当性の高い尺度であり、睡眠の質・量的情報が得られ、個体間や群間での比較が可能となる。

◆一般的信頼尺度[24.25]

他者をどの程度信頼しているかを表す心理指標である。一般的信頼尺度は、「ほとんどの人は基本的に正直である」「ほとんどの人は信用できる」「ほとんどの人は基本的に善良で親切である」「ほとんどの人は他人を信頼している」「私は人を信頼するほうである」「たいていの人は、人から信頼された場合、同じようにその相手を信頼する」という6つの質問に7件法で回答する。合計得点が高いほど人を信じる程度が高くなるとされる。本研究では総得点を用いた。

◆簡易嗅覚同定テスト（においスティック）（OSIT-J：The Odor Stick Identification Test for the Japanese）

嗅覚の同定能力（何のにおいがわかる能力）の検査に用いる簡便なキットであり、日本人に馴染みのある12種類の多様なにおいがマイクロカプセル化され、スティック型容器に納められている。各々のにおいについて、被検者は正解を含む4つの選択肢の中から嗅いだにおいに最もあてはまる選択肢を選び、4つの中に適当なものが

無い場合は「分からない」を、また無臭であると感じた場合は「無臭」を選択し、全 12 個中で正解した数でその被検者の同定能力を評価する。9/12 個以上の正解が正常の嗅覚同定能力を有すると示されているが、65 歳以上では、平均が 9/12 点との報告も認める。高齢者を対象とする本研究では、カットオフを 9 点とすると脱落が増えすぎてしまうため、6 点を選択基準におけるカットオフとした。

表2 評価項目のスケジュール

[illegible]

(7)分析手法

本研究においては、評価指標の計測は開始時、介入終了時の2時点で行われた。データ解析は、評価尺度結果について混合効果モデルによる反復測定データ解析法（Mixed Model Repeated Measures : MMRM）を行い、全参加者のデータを用いたITT（Intention-to-Treat）解析で行われた。各群の参加者の基本統計量の違いは、分散分析及びカイ二乗検定によって検証した。事後追加された二次解析は分散分析によって行われた。分析はプロトコルの作成に関与しない第3者が割り付けをマスクキングされた状態で実施された。分析の有意水準は5%とし、ブラインド解除前の一次的解析およびブラインド解除後の二次的解析とも「SPSS」が解析ソフトとして用いられた。検査結果が明らかに不自然な2症例を認め、それぞれに電話で直接確認したところ、両者とも真剣に検査に臨むことができなかったとの回答が得られたため、厚生省医薬品安全局通知「臨床試験のための統計的原則」に従い、研究者内で検討後、ITTとしても解析対象として不適切と判断されたそれらの2症例が、解析対象より除外された。

3. 結果

(1) 参加者の属性

参加者の属性を表 3 に示した。基準を満たし本研究の参加に同意した対象者 119 人が登録され、2 群に無作為割付がなされた。

ベースラインにおける平均年齢は両群とも 69.5 歳で、男性の割合はアロマ群で 53.3% (32/60 人)、プラセボ群で 45.8% (27/59 人) であった。また、平均 MMSE の平均得点は、アロマ群で 28.5 [±1.59] 点、プラセボ群では 28.6 点 [±1.25] 点であった。これらは、それぞれ有意な群間差は認めなかった。また、においスティック正解数でも両群間において有意差は認めなかった。他の属性については、過去の喫煙習慣に有意差を認めた ($p=0.0018$) 以外は、群間に有意差は認めなかった。

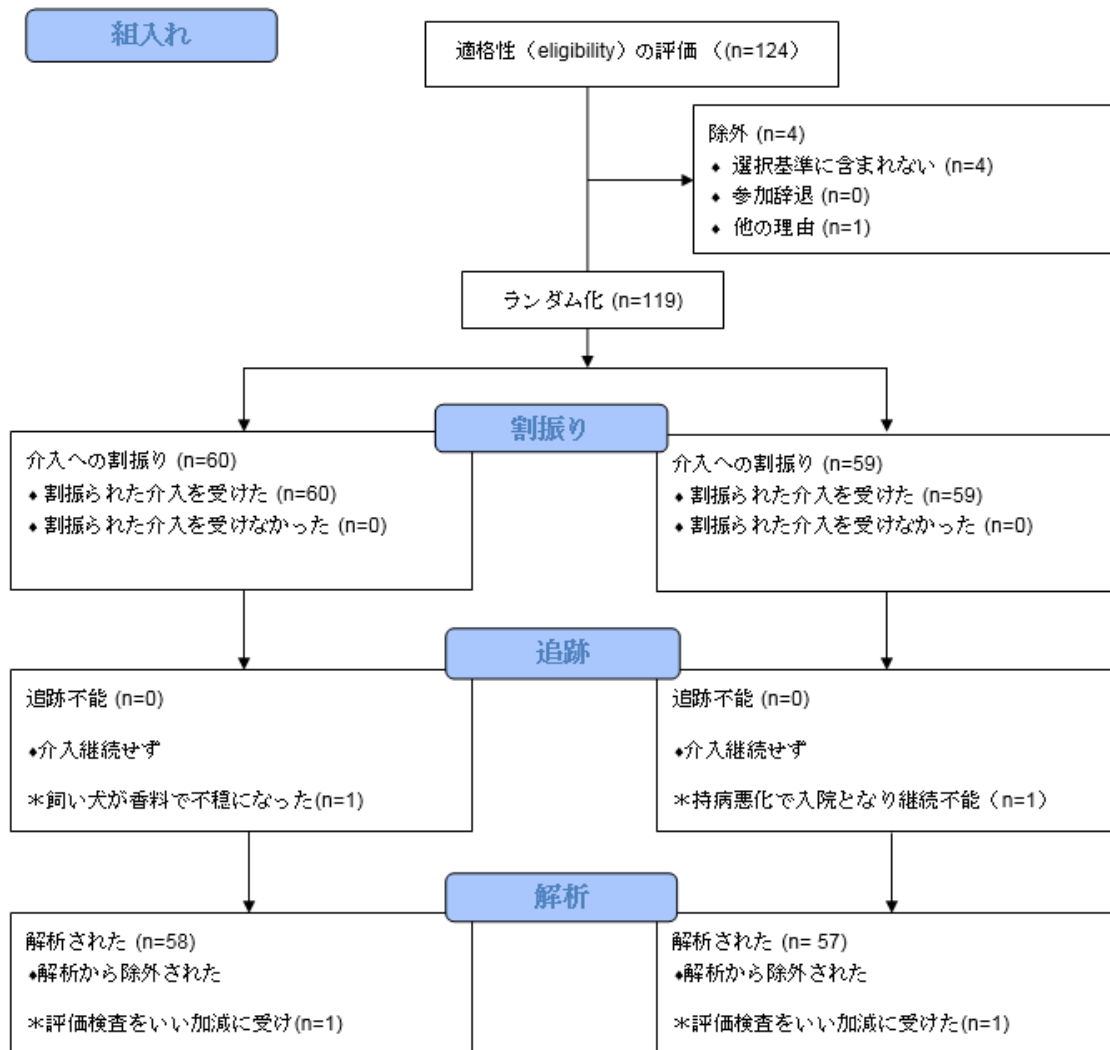
また、ベースラインにおける、PASAT の 2 秒条件および 1 秒条件、RAVLT (想起数合計・遅延再生・再認)、SDMT、PHQ-9、GAD-7、CES-D のウェルビーイング該当 4 項目の総得点と各項目点、SDS の 3 尺度、ピッツバーグ睡眠質問票の 7 項目および総合得点、一般信頼尺度についても、それぞれ群間に有意差は認めなかった (表 4)。

本研究はランダム化比較試験のため、喫煙習慣を共変量とした調整は行われずに解析は実施されたが、事後に喫煙習慣についてのみ主要評価項目に対して調整後の結果との比較確認を通じて感度分析が実施された。

(2) アドヒアランス

アロマ群で 2 名が辞退したがその理由は、1 人は自身の持病治療で入院となったから、もう 1 人は精油の香りに飼い犬が反応して不穏になったからという有害事象によるものであった。実施に関するアドヒアランス (実施率) では、傾向差でアロマ群が低かったが両群とも 96% 以上の参加者で実施率 80% 以上を達成が報告されていた。フローダイアグラムは図 5 に示した通りであった。

図5 CONSORT 2010 フローダイアグラム



(3) 評価項目

[1] 主要評価項目

聴覚性注意機能の評価検査である PASAT では、2 秒条件においては介入による変化は、アロマ群における正答数の得点増加が平均 5.80 (95%CI [3.76 - 7.84]) に対して、プラセボ群では平均 2.48 (95%CI [0.48 - 4.47]) とその平均差は 3.32 (95%CI [0.47 - 6.17]) と有意な ($p=0.0223$) 改善を認めていた (図 6) 。得られた効果量は、Hedges' $g=0.44$ であった。

一方で、より難度の高い 1 秒条件においては、アロマ群における正答数の得点増加が平均 1.06 (95%CI [-0.81 - 2.93]) に対して、プラセボ群では平均 2.24 (95%CI [0.41 - 4.07]) ($p=0.374$) と、介入前後の時系列推移 (= 変化) に有意な群間差は認めなかった (図 7) 。

図6 健常高齢者の注意機能に対するアロマセラピーの効果 (PASAT-2 秒条件)

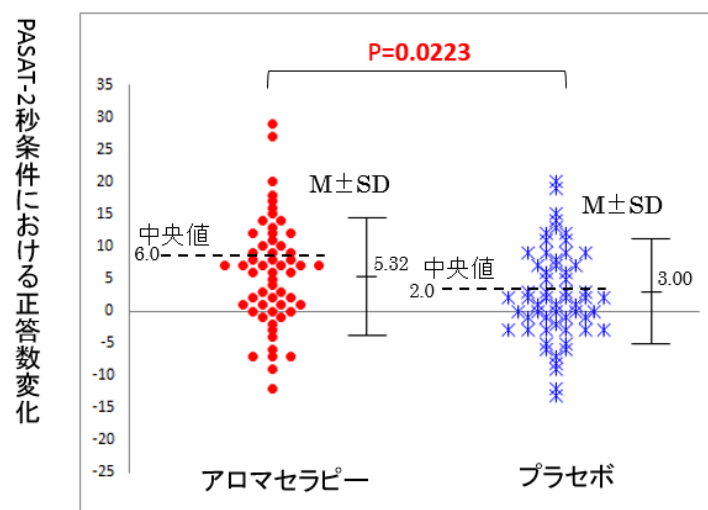
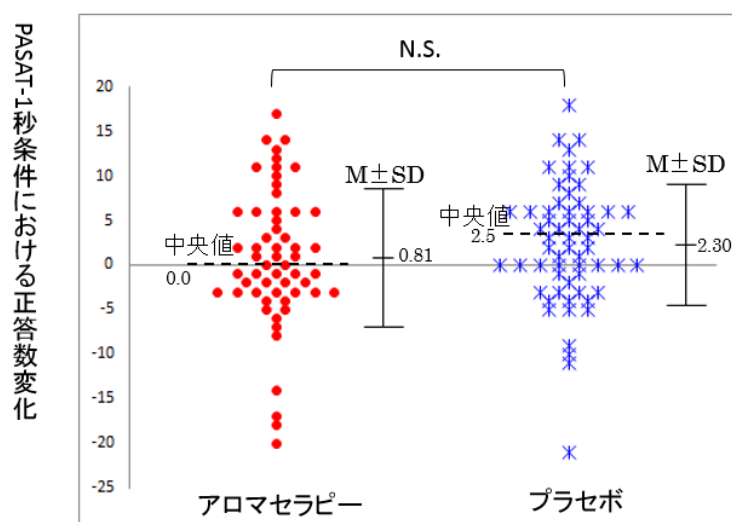


図7 健常高齢者の注意機能に対するアロマセラピーの効果 (PASAT-1 秒条件)



[2] 副次的評価項目

①RAVLT：想起数合計、遅延再生（干渉後）、再認

想起数合計である全即時記憶（total immediate memory :TIM）では、アロマ群における正答数の得点増加は平均 1.64（95%CI [0.80 - 4.09]）に対して、プラセボ群では平均 3.95（95%CI [1.51 - 6.39]）（ $p=0.189$ ）と、介入前後の時系列推移に有意な群間差は認めなかった。

学習能力の指標となる達成想起数（第 5 試行-第 1 試行）である言語学習能力（verbal learning ability : VLA）では、アロマ群においては前後変化で平均-1.09（95%CI

[-1.75 - 0.42]) に対して、プラセボ群では平均-0.68 (95%CI [-1.34 - -0.01]) (p=0.394) と、両群とも介入前後の時系列推移に有意な群間差は認めなかった。

次いで学習能力の指標となる第5試行での想起数では、アロマ群においては前後変化で平均 0.32 (95%CI [-1.15 - 0.80]) に対して、プラセボ群では平均-0.37 (95%CI [-0.09 - 0.84]) (p=0.881) と、介入前後の時系列推移に有意な群間差は認めなかった。

認知症の記憶能評価では最も鋭敏で重要な近似記憶指標とされる遅延再生（干渉後）においては、アロマ群における正答数の得点増加が平均 0.56 (95%CI [-0.12 - 1.23]) に対して、プラセボ群では平均 0.67 (95%CI [0.00 - 1.34]) (p=0.811) と、介入前後の時系列推移に有意な群間差は認めなかった。

再認では、アロマ群における正答数の得点増加は平均-0.30 (95%CI [-0.62 - 0.03]) に対して、プラセボ群では平均 0.15 (95%CI [-0.17 - 0.47]) (p=0.054) と、介入前後の時系列推移に有意な群間差は認めなかった。

②SDMT :

アロマ群における正答数の得点増加は平均 0.86 (95%CI [-0.42 - 2.15]) に対して、プラセボ群では平均 0.39 (95%CI [-0.80 - 1.57]) (p=0.589) と、介入前後の時系列推移に有意な群間差は認めなかった。

③MMSE :

MMSE では、遅延再生（干渉後）においては、アロマ群における正答数の得点増加は平均-0.08 (95%CI [-0.63 - 0.48]) に対して、プラセボ群では平均 0.28 (95%CI [-0.28 - 0.84]) (p=0.375) と、介入前後の時系列推移に有意な群間差は認めなかった。

④PHQ-9 :

PHQ-9 では、アロマ群における介入前後での得点変化は平均 0.32 (95%CI [-0.29 - 0.92]) に対して、プラセボ群では平均 0.37 (95%CI [-0.23 - 0.97]) (p=0.905) と、介入前後の時系列推移に有意な群間差は認めなかった。

⑤GAD-7 :

GAD-7 では、アロマ群における介入前後での得点変化は平均-0.22 (95%CI [-0.71 - 0.27]) に対して、プラセボ群では平均-0.15 (95%CI [-0.64 - 0.34]) (p=0.850) と、介入前後の時系列推移に有意な群間差は認めなかった。

⑥SDS :

i) 仕事／学業

SDS_仕事／学業では、アロマ群における介入前後での得点変化は平均 0.06 (95%CI [-0.14 - 0.25]) に対して、プラセボ群では平均 0.15 (95%CI [-0.04 - 0.34]) (p=0.497) と、介入前後の時系列推移に有意な群間差は認めなかった。

ii) 社会生活

SDS_社会生活では、アロマ群における介入前後での得点変化は平均-0.05 (95%CI [-0.30 - 0.21]) に対して、プラセボ群では平均 0.12 (95%CI [-0.13 - 0.37]) (p=0.366) と、介入前後の時系列推移に有意な群間差は認めなかった。

iii) 家庭内のコミュニケーションや役割

SDS_家庭内のコミュニケーションや役割では、アロマ群における介入前後での得点変化は平均-0.10 (95%CI [-0.37 - 0.17]) に対して、プラセボ群では平均 0.13 (95%CI [-0.13 - 0.40]) (p=0.226) と、介入前後の時系列推移に有意な群間差は認めなかった。

⑦ウェルビーイング (CES-D 該当 4 項目 : Q4、8、12、16)

ウェルビーイングでは、アロマ群における介入前後での得点変化は平均-0.19 (95%CI [-0.95 - 0.56]) に対して、プラセボ群では平均 0.93 (95%CI [0.19 - 1.68]) であり、その平均差は-1.13 (95%CI [-2.19 - -0.06]) とプラセボ群が有意に勝つという結果を認めた (p=0.038) 。

⑧一般的信頼尺度

一般的信頼尺度では、アロマ群における介入前後での得点変化は平均 0.28 (95%CI [-0.93 - 1.50]) に対して、プラセボ群では平均 0.53 (95%CI [-0.67 - 1.73]) (p=0.772) と、介入前後の時系列推移に有意な群間差は認めなかった。

⑨ピッツバーグ睡眠調査票

C1 睡眠の質

睡眠の質では、アロマ群における介入前後での得点変化は平均-0.06 (95%CI [-0.20 - 0.08]) に対して、プラセボ群では平均-0.05 (95%CI [-0.19 - 0.09]) (p=0.917) と、介入前後の時系列推移に有意な群間差は認めなかった。

C2 入眠時間

入眠時間では、アロマ群における介入前後での得点変化は平均-0.08 (95%CI [-0.23 - 0.08]) に対して、プラセボ群では平均-0.05 (95%CI [-0.20 - 0.10]) (p=0.788) と、介入前後の時系列推移に有意な群間差は認めなかった。

C3 睡眠時間

睡眠時間では、アロマ群における介入前後での得点変化は平均 0.01 (95%CI [-0.15 - 0.17]) に対して、プラセボ群では平均 0.03 (95%CI [-0.12 - 0.19]) (p=0.860) と、介入前後の時系列推移に有意な群間差は認めなかった。

C4 睡眠効率

睡眠効率では、アロマ群における介入前後での得点変化は平均-0.06 (95%CI [-0.18 - 0.07]) に対して、プラセボ群では平均 0.08 (95%CI [-0.05 - 0.20]) (p=0.150) と、介入前後の時系列推移に有意な群間差は認めなかった。

C5 睡眠困難

睡眠困難では、アロマ群における介入前後での得点変化は平均 0.06 (95%CI [-0.10 - 0.21]) に対して、プラセボ群では平均 0.02 (95%CI [-0.14 - 0.17]) (p=0.711) と、介入前後の時系列推移に有意な群間差は認めなかった。

C6 眠剤の使用

眠剤の使用では、アロマ群における介入前後での得点変化は平均 0.02 (95%CI [-0.08 - 0.12]) に対して、プラセボ群では平均-0.02 (95%CI [-0.12 - 0.08]) (p=0.629) と、介入前後の時系列推移に有意な群間差は認められなかった。

C7 日中覚醒困難

日中覚醒困難では、アロマ群における介入前後での得点変化は平均 0.14 (95%CI [0.00 - 0.29]) に対して、プラセボ群では平均 0.00 (95%CI [-0.14 - 0.14]) (p=0.174) と、介入前後の時系列推移に有意な群間差は認めなかった。

□ピッツバーグ睡眠質問票総合得点 (PSQIG)

睡眠総合得点では、アロマ群における介入前後での得点変化は平均-0.01 (95%CI [-0.48 - 0.46]) に対して、プラセボ群では平均 0.02 (95%CI [-0.45 - 0.48]) (p=0.932) と、介入前後の時系列推移に有意な群間差は認めなかった。

⑩簡易嗅覚同定検査 (においスティック)

簡易嗅覚同定検査 (においスティック) では、アロマ群における介入前後での得点変化は、アロマ群では平均-0.54 (95%CI [-0.95 - -0.12]) に対して、プラセボ群では平均 0.22 (95%CI [-0.19 - 0.62]) であり、その平均差は-0.75 (95%CI [-1.34 - -0.17]) とプラセボ群が有意に勝るという結果を認めていた (p=0.012) 。

4. 考察

本研究は、我々の知る限り健常高齢者に対するアロマセラピーによる認知機能改善を厳密なランダム化比較試験で示したはじめての報告であった。脱落は 4 名と低く、アドヒアランスも両群とも 96%以上の参加者において 80%以上の実施率達成で報告されるなど高く、参加者全体の参加モチベーションも終了時アンケートで 1 (低い) -10 (高い) の 10 段階評価で平均 8.46 [SD=1.63] と高いものであった。アロマセラピー群に対してアクティブ・プラセボとしてのエタノール群を設定し、参加者は全員が簡易嗅覚同定検査で半数 (6 個) 以上の正答可能な最低限の嗅覚機能を有していた上 (参加者平均でも 3/4 にあたる 9 個を正答)、計 4 時間/日以上精油による嗅覚曝露が、高いアドヒアランスで行われ、ランダム化比較試験 (RCT) として高水準で実施されたと考えられた。

以下、1) 認知機能への効果、2) 心理社会的側面への効果、3) 嗅覚への効果、4) アロマセラピーにおけるランダム化比較試験という研究方法の妥当性について、それぞれ考察を加えたい。

1) 認知機能への効果

平均年齢 69.5 歳であった本研究でのベースラインにおける PASAT-2 秒条件の平均正答率は 63.3%で、SDMT 平均達成率は 47.3%と、先行研究で示されている日本人健常高齢者との比較ではほぼ平均と同等水準であった。また、MMSE のベースラインにおける平均得点も 28.4 (SD=2.80) は、同年代の地域調査での 70-79 歳の 28.0 (SD=1.9) とほぼ同じであり、RAVLT のベースラインにおける平均の総想起数 48.0 (SD=9.63) ・遅延再生 (干渉後) 数 9.93 (SD=3.05) ・再認数 14.4 (SD=1.11) は、やはり同調査の 49.5 (SD= 8.1) ・ 9.8 (2.8) ・ 13.8 (1.2) とほぼ同等の結果であった。以上より、本調査の標本代表性には問題がなかったと言える。

介入効果に関しては、主要評価項目である PASAT のうち 2 秒条件においては、アロマ群がプラセボ群に対して有意差を認めた。一方で PASAT の 1 秒条件、および副次的評価項目である SDMT、RAVLT (想起数合計・遅延再生 (干渉後) ・再認)、MMSE のすべてにおいてアロマ群とプラセボ群の間に有意差は認められなかった。主要評価項目の PASAT では、2 秒条件と 1 秒条件の 2 項目が設定されていたため、厳密には既存の多くの介入研究の報告では検討されずに無視されることも多いが、統計学的な多重検定の危険性も生じている[25]。

これは、評価対象が増えるほど偶然の産物として有意差を認めたかのような偽陽性を生じる危険性のことであり、本来は統計学的補正を要する。本研究では主要評価項目に対してボンフェローニ法を実施した場合、2 秒条件と 1 秒条件の両者が無相関との前提でも p 値である 0.0223 を 2 倍した $p=0.0446$ が有意水準である 5% (0.05) を下回るために、PASAT-2 秒条件で示された有意差の信頼性は高いと言える。実際には 1 秒条件と 2 秒条件は、強い相関性を有しており (相関係数=0.73)、補正後 p 値は 0.0446 よりも小さな値になると想定されるが、本結果が有意であるという結論が変わらないこともあり本稿ではそこまでの算出は割愛した。

2020 年 8 月に更新された最新のコクラン・レビュー (2020) [26]でも、これまで実施された 3 つランダム化比較試験のすべてにおいて認知症患者に対するアロマセラピーによる認知機能の有意な改善は依然認められてはいなかったことが示されている。また、健常高齢者を対象としたランダム化比較試験において、アロマセラピーによって中核的認知機能の改善が厳密に示されたという報告は我々の知る限りこれまで皆無である。そのため本研究で得られたアロマセラピーにより認知機能改善が示されたという結果の持つ意義は大きいと考えられた。

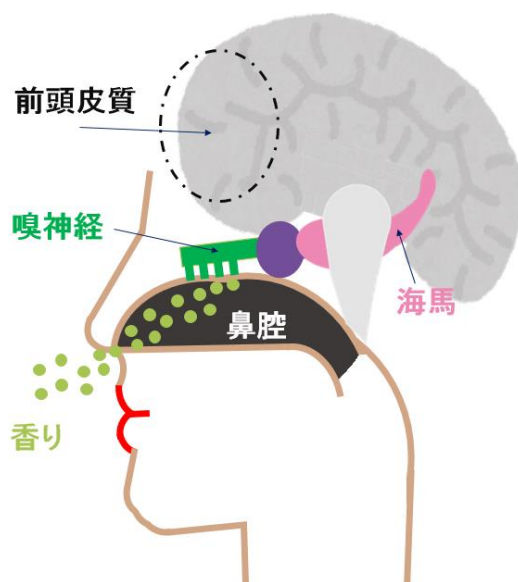
なぜ、2 秒条件だけに効果が示されて 1 秒条件では示されなかったかについては、2 秒条件では 2 秒で行う計算課題を、1 秒条件では 1 秒という半分の短時間で行うことが要求されるために作業負荷がかなり大きい。そのため、課題としてより難易度の高い 1 秒条件では差が認められなかったのに対して、同じ課題を倍の時間を費やすことが許され、相対的に容易な課題である 2 秒条件だけにおいてアロマセラピーの効果が評価に反映されやすいということは、合理的説明がつくとは考えられた。

PASAT の基本的な評価標的は「注意」である。この改善が認められる時には、その前提として注意の連続的な変換能力である分配性注意が反映しているとされる。一方

で同じ「注意」を評価標的とする SDMT は、PASAT が聴覚的注意なのに対して、視覚的注意の評価とされている。ふたつの注意評価テストでアロマセラピーによる差が出た背景にはこの感覚モダリティの違いが影響している可能性があるのかもしれない。しかし、それ以上に PASAT は SDMT に比べて、ワーキングメモリ（作業記憶）への負荷がより大きな検査とされており、注意改善の背景にはワーキングメモリへの芳香による嗅覚刺激への働きかけが寄与していた可能性が示唆された[27]。ワーキングメモリは、現在の作業に必要な情報を一時的に記憶し、その記憶に基づいて一連の作業を効率的に実行するための機能である。Chamine らは、平均 50 代後半の健常者に肉体的（氷水に手をつける）、感情的（不愉快なスライドを見せられる）、知的（モントリオールイメージ課題：制限時間内で計算を行う課題）という複数のストレス課題を実施した後、ラベンダー吸入群と対照群とに分けるランダム化比較試験によって、ワーキングメモリ負荷のかかる認知試験においてラベンダー吸入群が対照群に 15%ほど有意に優れていたとの報告をしている[28]。この結果は、ラベンダーが成分を含む当研究での介入による PASAT 改善が、ワーキングメモリ改善への寄与を介していた可能性を示唆するものと言える。

ワーキングメモリは、これまで前頭皮質を中心とした脳領域の関与が注目されてきたが近年ラットなどでは海馬やその周辺領域などの大脳辺縁系の関与を示す報告がなされてきている[29]。アロマセラピーは、嗅覚器から直接大脳辺縁系に作用する神経学的伝達経路を有し、より直接的に働きかけることから、ワーキングメモリ改善の機序としてアロマセラピーによる海馬領域の活性化を通じて寄与した可能性はあるとも考えられた（図8）。

図8 香りが嗅覚から脳に伝わるメカニズム



注意は認知機能の一つであり、大脳に届く膨大な量の情報の中で必要な一部分だけを適時適切に取捨選択するというものである。言い換えるならば「大切なことに意識を集中する働き」である。我々は、今回の主要評価項目を“記憶”よりも“注意”の改善に主眼をおいた。

その理由のひとつとして、記憶に問題のない健常高齢者である参加者を募った本研究では記憶の改善効果は天井効果の観点から難しいと想定されたことが挙げられる。一方で、多くの先行研究により注意は遂行機能（≡問題解決能力）と共に MMSE 正常の健常高齢者であっても年齢を経る毎に、脳機能の中でも記憶などの他機能よりもまず先に低下がみられ、またそれを統括する前頭葉機能の低下も示唆されるという現象が、前頭葉加齢仮説（frontal aging hypothesis）として知られている[30]。そして注意障害は認知症においても早期から出現することが知られ[31]、認知症予備軍の MCI 段階でも分配性注意の障害は顕著に認められることも報告されている[32]。何より、注意はあらゆる認知機能の根幹であり、注意が障害されるとその上位にあたる知覚や思考、記憶や学習、遂行機能といった高次脳機能のすべてが毀損される[33]。そのため、認知症における認知障害の理解にも必須な要素と言える。それにも関わらず、現実には特に記憶ほどには、注意に対してこれまで十分には注目されてきたとは言えなかった[34]。

しかし近年は、初期の認知機能の低下状態では、注意は記憶以上に臨床的に重要な意味を有するとの議論もなされてきている。一般に認知症というと記憶低下というイメージが強い。勿論、記憶障害は進行した認知症における重要な中核的症状であることには疑いないが、近年はアルツハイマー型認知症での記憶障害の陰に潜む、注意障害こそが生活管理能力の低下において、記憶障害よりも大きな影響を与えている可能性が示されている。「注意」は、高次脳機能の中でも他の認知過程の土台となり、その障害は多くの日常・社会生活を阻害しうる[35]。

例えば、アルツハイマー型認知症では初期から金銭や服薬の管理、移動、家事などの自立した生活能力の指標ともいえるべき手段的日常生活動作（IADL : instrumental activities of daily living）が低下する。船山らによれば[36]、軽度の認知障害（MCI）やアルツハイマー型認知症の患者に対して行った調査で、IADL への影響因子を、1) 知能（改訂長谷川式知能評価スケール）、2) 見当識（MMSE の該当箇所 10 点分）、3) 注意（ウェクスラー成人知能検査 3 版（WAIS-III）の符号問題）、4) 出来事記憶（ウェクスラー記憶検査改訂版（WMS-R）の論理記憶）、5) 遂行機能（前頭葉機能検査（FAB））の因子での解析の結果、有意に影響するのは注意障害のみであったと報告する。つまり、初期の認知機能低下者の自立的な生活の破綻は、記憶力の低下や見当識障害によってではなく注意機能の低下による可能性が高いということである。近年、信号無視やアクセルとブレーキの踏み間違いといった高齢者の自動車の危険運転の予測因子としても、神経心理学的評価としては「記憶障害」は含まれておらず、「注意障害（注意の分配）」が重視されている[37]。これは、アルツハイマー型認知症に限らず、レビー小体型認知症などでも同様のことが議論されている。

PASAT-2 秒条件で得られた効果量は、Hedges' $g = 0.44$ であった[38]。これは、Cohen の提唱した「小（0.2）、中（0.5）、大（0.8）」という基準に従えば、中等度相当の効果量と解釈される[39]。しかし、Thompson らは、このような効果量の規範に基づいて画一的に解釈される現在の慣習に異を唱え、効果量は先行研究での文脈に基づいて解釈される必要性を主張する[40]。本効果量は、わかりやすく「偏差値」に例えるならば、3 か月のアロマセラピーで、注意能力においてライバルに偏差値で 4 以上の差をつけることができた、とも言い換えられる。本研究の対象者は健常高齢者（将来的に

は認知機能低下者へのアロマセラピー応用も念頭に置くが）であり、その目的はこれから劇的に注意能力を高めることではなく、楽しみとしてのアロマセラピーを通じて無理なくその低下を予防することで自立した現状生活の維持や、生活上のリスク低減にあるという視点に立つならば、本結果の臨床的意義は得られた効果量以上とも解釈できるかもしれない。

尚、近年はランダム化比較試験の場合、ベースラインにおける属性等において有意差が認められても共変量調整は行われずに主解析が実施される流れにあり、本研究もそれに倣った。しかし、事後に主要評価項目に対して喫煙習慣の調整後結果も算出し、結果に違いがないことを感度分析として比較確認している。

注意を主として評価対象とする PASAT や SDMT に対して、RAVLT は記憶（言語性の近時記憶）や学習能力に関する評価における代表的検査である。神経心理学的に、近時記憶は情報の記録後すぐに想起させる即時記憶より保持時間の長い記憶であり、記録と想起の間に他の刺激や質問による干渉が介在されて保持情報が一旦意識から消えても再生可能な記憶という特徴を有する。アルツハイマー型認知症では、特に初期にはこの近時記憶の障害がより顕著であることが知られる。実際、アルツハイマー型認知症に進みやすい MCI の記憶障害の特徴は、やはり新しい情報の学習能力や近時記憶の障害である[41]。

RAVLT は同じ単語のリストを第 1 試行から第 5 試行まで繰り返し記憶し、干渉課題をはさんだ後に、思い出せる数を測定（遅延再生）する単純な検査であるが、エピソード記憶をよく反映する[42]。通常①学習能力は、第 5 試行での想起達成数から第 1 試行の達成数から第 1 試行の達成数を引いた言語学習能力（verbal learning ability : VLA）や、第 5 試行における想起数などから評価される。また、その後の干渉課題を経たからの遅延後の記憶の保持（近時記憶）を、遅延再生（干渉後）と再認（見覚えがあるかの確認）で測定することで、アルツハイマー型認知症でも重要な②記憶機能を分析できる[43]。Eckerström ら（2013）は MCI からアルツハイマー型認知症へ進展することを予測する因子として他の神経心理学的検査、脳脊髄液の生物学的指標、MRI での海馬容積などと比較しても単独でのベストは RAVLT であったと報告するなど優れた記憶検査であることが知られている[24]。

結果として、本研究では記憶に対するアロマセラピーの効果については、RAVLT において VLA（第 5 試行－第 1 試行）、第 5 試行想起数、総想起数（第 1 試行から第 5 試行の達成数の総計である全即時記憶（total immediate memory :TIM）、遅延再生（干渉後）、再認のすべてにおいて有意差は認められなかった。アルツハイマー型認知症に転換しやすい MCI の記憶障害の特徴として最も関心の高い評価項目とされている近似記憶としての遅延再生に[41]変化が認められなかった上、再認でも差を認めなかったことはその結論を支持する。また、学習能力の反映である想起数でもやはり差を認めなかったことから、本研究においてはアロマセラピーが近似記憶や学習能力を含めた“記憶”について有意な影響を与えているとの結論には至らないと考えられた。

その説明として、先行研究からは RAVLT においては MMSE 正常（開始時平均得点 28.5 点）であり、物忘れで困っていないという条件でリクルートされた、健常高齢者が対象であったことが考えられる。健常高齢者においては天井効果（学習実験などで

試行を重ねても遂行がそれ以上伸びない頭打ち状態となる現象）から効果が得られなかった可能性は示唆された。

包括的な認知機能スクリーニング検査である MMSE でも、アロマセラピーには有意な差は認められなかった。この結果もそもそも MMSE 正常者を対象としていた上、開始時の平均が 28.5 点と中でも高得点であったことから、やはり天井効果により改善の余地は殆どなかったためと考えられる。

アロマセラピーの認知機能への効果について総括すると、本研究における健常高齢者に対するアロマセラピーによる効果は、アルツハイマー型認知症の中核的病理である学習能力や近時記憶に対しての効果は認められなかった一方で、認知症だけに限らない高齢者の生活リスクを生じさせる最大の要因とも考えられる注意機能（特に、聴覚的な注意の分配に関して）に対しては、有意な効果が示されていたということになる。

2) 心理社会的側面への効果

抑うつ指標 PHQ-9 や不安指標 GAD-7 においては、総得点において有意な変化は認められなかった。これは、そもそもベースラインにおける平均得点がうつ病や全般不安症の閾値下の正常であり、天井効果によって両群に差が生じなかったことは妥当な結果と解釈された。

抗うつ効果については、2017 年に 12 のランダム化比較試験についての系統的レビューが示されている[44]。そのうち吸入だけの抗うつ効果を評価した研究は 5 つであったが、有意な抗うつ効果が示されたのは 2 研究のみであり、残りの 3 研究は否定的な結果に至っている。それに対して、アロママッサージで実施された 8 研究では、5 研究で有意な効果が示されている一方で、それらの結果がアロマセラピーの香りによるものなのか、マッサージによるものなのかは不明瞭であるとされている。現状ではマッサージそのものによる抗うつ効果も系統的レビューで示されており[45]、どの程度香りによる要素が抗うつ効果に寄与しているのかはさらなる検証を要する[46]。

同様に、抗不安効果についても Lee らの 16 研究によるメタ解析では、7 研究が吸入で、6 研究はマッサージで実施されていた。抗うつ効果と同様にアロマ吸入だけによる効果か、マッサージの効果によるものなのかは議論の対象となるが、不安に対しては 4 研究で、アロマオイルを使用したマッサージと、アロマオイルを含まないプラセボオイルによるマッサージの比較が検証され、3 研究でアロマオイルがあることでより効果が高まることが傾向差で示されている。

しかし、アロマセラピーによる抗うつ効果や抗不安効果の研究は、ランダム化比較試験で行われていたとしても研究デザインの質が高いとは言えず、また研究間の異質性も高いため結論が出せる水準になく、今後さらに良質の検証を重ねる必要がある。

ウェルビーイングに関しては、先行研究では身体疾患を有する患者に対してのウェルビーイングを高めるとの報告も複数認められていたが、本研究はその結果には反するも

のであった。郡内での前後比較を見てみると、アロマ群では平均で-0.19 (95%CI [-0.95 - 0.56]) と前後で有意な変化を示していないのにも関わらず、プラセボ群では0.93 (95%CI [0.19 - 1.68]) と有意な改善が認められていた。このため何かしらの理由でアロマセラピーがウェルビーイングを悪化させたというよりは、プラセボ群が有意に改善したために群間での有意差が生じたと解釈される。たとえば、プラセボのエタノールによる酩酊効果で多幸等が生じた可能性も浮かび上がる。しかし、1回使用量が極端に少なく、揮発性の高さからは嗅覚刺激効果の持続時間も短いと考えられるエタノールの吸入ではその可能性は考えづらい。これは例えば、日々高濃度のアルコール綿を使って消毒作業を行う採血業務に従事する看護師が、酩酊や多幸には至らないのと同じ理由で、“エタノール吸入幸福説”への帰属は否定的に考えざるを得ない。この原因は本研究のみでははっきりとはしない。介入内容の異質性から同列には扱えないが、似たような対象者として嗅覚障害を有する高齢者に対するランダム化比較試験で、朝晩2回/日×5か月にわたる嗅覚トレーニングによって嗅覚や抑うつは有意に改善したが、ウェルビーイングで改善を認められていなかった[20]。

なお、事後の探索的な追加解析として、通常はアロマセラピーで改善されると信じられているウェルビーイングにおいてネガティブな方向に働いている背景に、嗜好性を問わずに一律の精油を用いている研究デザインにあるのではないかと考え、使用した精油の香りの好き嫌いを上下で二群したものを共変量として加えて、MMRMで再解析を行った。結果、アロマセラピーに望ましくないという結果において有意差が認められなくなった。一方で、調整後にもアロマセラピーに有利な結果が認められることもなかった。そのため、個人の嗜好に合わない香りがウェルビーイング悪化に寄与した可能性が一定程度は示唆されるかもしれない。アロマセラピーが日本に入ってきたのは1980年代であり、日本で盛んにアロマセラピーが行われるようになったのは1990年代後半と言われる。つまり大衆化したのはここ20年であり、65～80歳の方々のアロマセラピー経験値は低い。またブレンドされたアロマオイルの場合「何の香りかわからない」ことで、混乱と警戒によりネガティブな感情があったことも考えられる。

また、本ウェルビーイングは厳密には15項目以上設定された副次的評価のひとつであり（睡眠はひとつの質問として扱ったとして）、統計学的多重性を考慮してボンフェローニ補正を実施した場合、 $p=0.032$ で示された有意差は消失する。一度に15項目を多重測定した場合、そのような研究では誤って偶然に有意差が生じる偽陽性の可能性は少なくとも50%以上とも示されている[25]。上記考察は、その前提としてあくまでも本研究は探索的位置づけの上での副次的評価という前提の上での議論であり、たとえ有意差があっても決定的結論には至らないということは念頭に置かなければならない。

睡眠に関しても、本研究ではアロマセラピーのよって改善が得られるという結果には至らなかった。Linらによる31のRCT研究を対象とした最新のメタ解析によっても[47]、アロマセラピーは睡眠の質を有意に改善することが、Hedges' $g=1.10$ (95%CI [0.81 -1.39]) という高い効果量で示されている。このメタ解析では、13研究は本研究と同様に健常者を対象としたものであった。また、吸入よりもマッサージに分のあった抗うつ効果や抗不安効果と異なり、睡眠では吸入のみでも非吸入よりも有意に高い

睡眠効果も示されており、複数のアロマオイル混合においても有意な効果量が示されていることから、本研究でも効果は期待されるものであった。また対象年齢も平均年齢で 60 代以上が半数近く含まれており、メタ回帰分析でも年齢については効果に違いがないことも示されていた。本研究で採用したピッツバーグ睡眠調査票は、Lin らの報告でも過去の研究では最頻用されている調査票であったが、総合評価だけでなく種々の関連する下位項目（睡眠の質、入眠時間、睡眠時間、睡眠効率、睡眠困難、眠剤の使用、日中覚醒困難）にわたって各々について評価がなされるが、それらのすべてにおいて本研究では群間差が認められなかった。

アロマ群に効果はあったが、プラセボ群のエタノールにも効果があったために差が相殺されて認められなかったという可能性も、それぞれの群内での前後比較で有意差が認められなかったことから否定的で、逆にどちらも睡眠に対しては益も害もなかったと示唆された。

3) 嗅覚同定機能への効果

本研究では、アロマ群がプラセボ群に対して、有意な嗅覚同定能の低下が示されていた。群内変化として、プラセボ群では前後変化が認められなかったのに対して、アロマ群では平均-0.54 (95%CI [-0.95 - -0.12]) 点と軽度な低下が有意に認められ、このため群間の差は、アロマセラピーによる嗅覚低下に起因すると考えられた。

嗅覚障害を有する患者でもいわゆる嗅覚トレーニングで改善が得られる報告が複数なされてきている[48-51]。また高齢者の嗅覚低下に対しても、朝晩の嗅覚刺激で有意な改善が得られたとの Birte-Antina らの報告もある[20]。しかし、本研究ではそれらの先行研究に一致した結果が得られなかった。高齢者の嗅覚障害に有意な改善が示されたとはいっても、あくまでも改善した高齢者の嗅覚は嗅覚閾値と嗅覚識別能力だけであり、本研究で評価対象とした嗅覚同定能力までの改善を示す RCT は、我々の知る限り報告されていない。

効果が得られなかった原因としては、漂っている香りを自然呼吸で吸い込むアロマセラピーと、香りを能動的に嗅ぐ嗅覚トレーニングでは、気流やにおいの感じ方も異なることが知られており、嗅覚改善が得られなかったのはその影響が関係していた可能性があるかもしれない。また、先行研究では抑うつ状態は嗅覚低下をきたすという報告が示されているが、本研究はうつ状態では群間で有意差は認めていない。

さらに嗅覚悪化の要因として、今回採用された精油の曝露時間や頻度といった介入としての強度が強すぎたという可能性はありうる（アロマセラピーとしては一般的な範疇だが）。嗅覚には、いわゆる嗅覚疲労という現象が知られる。今回は健常とはいえ嗅覚機能の低下した高齢者に対して認知機能の改善を目指すという狙いもあり高強度の介入でデザインされていた。嗅覚刺激トレーニングによる嗅覚改善は、たとえば、一般に朝晩 10 秒ずつという短時間で 4 種類の香りを能動的に嗅がせる方法が、嗅覚障害の改善や嗅覚低下の予防になる、という報告は多数出ているが、本研究では曝露 1 回あたりの時間が長すぎた可能性は考えられる。あくまでも能動的な嗅覚訓練というよりは受動的なアロマセラピーであり、嗅覚訓練を 2 時間実施したわけではない

ものの、それでも嗅覚疲労に至った可能性は否定できない。しかし、高強度とはいっても、本研究程度の精油使用量はアロマセラピーとして一般的であり、かつ、この程度の強度でなければ認知機能改善は得られなかった可能性は否定しきれない。本研究だけでは、その原因特定までは難しいが、嗅覚改善を最大化する、もしくは嗅覚疲労を避けながら、認知機能改善や心理的効果を最大化するための至適強度及び頻度が、医科学的に十分に検討されているとは言えず、将来的に重要な検討課題のひとつではあると思われる。

また、ウェルビーイングでの議論と同様に、嗅覚機能も 15 項目以上設定された評価項目のひとつであり（睡眠はひとつの質問として扱ったとして）、統計学的多重性を考慮してボンフェローニ補正を実施した場合、 $p=0.012$ で示された有意差は消失する。一度に 15 項目を多重測定した場合、そのような研究では誤って偶然に有意差が生じる偽陽性の危険性は少なくとも 50%以上と示されている。そのため嗅覚機能の結果に関しても、あくまでも決定的結論には至らないという点を議論の前提として念頭に置いて解釈しなければならないということは踏まえるべきである。

4) アロマセラピーにおける集団対象でのランダム化比較試験という研究方法の妥当性について

本研究において特に、心理社会的な側面において、アロマセラピーに期待された効果が得られないといった先行研究に矛盾した結果が多く生じた背景要因のひとつとして、個人の嗜好や気分、体質を考慮せず一律に同じ香りを全参加者に、高強度に提供するという研究デザインが影響した可能性も推測された。

本研究では、事後解析として、にの「好き・嫌い」と「快・不快」についての評価では、朝用のにの液についてはアロマ群とプラセボ群で有意差は認めなかったが、夜用のにの液についてはアロマセラピー群がプラセボ群より「好き・嫌い」と「快・不快」において好き ($p=0.0877$)・快 ($p=0.0738$) との回答が有意傾向で多く回答されていた。シール貼付の面倒さや 3 か月間のモチベーションの有無についての意識調査では両群で有意差は認められなかった。また、「有益事象があった」との回答者は、精神的 ($P<0.001$)・身体的 ($p=0.0430$) のどちらにおいてもアロマセラピー群がプラセボ群に対して有意に多くの報告が認められた。一方で「有害事象があった」との回答者も、精神的 ($P=0.0140$)・身体的 ($p=0.0012$) のどちらにおいてもアロマセラピー群がプラセボ群に対して有意に多くの報告が認められた。

また、本研究の参加者は年齢相応ではあっても、嗅覚低下自体は加齢性に生じていたため、アロマセラピーの生物学的効果の最大化を狙って高強度での香りの曝露が時間・頻度ともに行われていたことも、香りの好き嫌いが悪影響を生じていた可能性はありえたかもしれない。例えば、低強度での介入あれば、特に心理社会的側面は改善が得られた可能性もありえたかもしれない。本来、嗜好性や体質を考慮すべきアロマセラピーに対して、一律の精油を用いて行われた現状の集団対象でのランダム化比較試験の枠組みに従った介入の方法論における限界を見た可能性も考えうる。

これは、他人には好ましい感覚であっても自分には嫌いな感覚は苦痛でしかないとすることは、聴覚、味覚、視覚、触覚を含む生理的なすべての他感覚にも共通するものであり了解可能である。特にアロマセラピーでは、自分で心地良いと感じる精油で受けることで、リラクセーションが促進され、睡眠障害等の改善に効果を認めると言われている。嗅覚は、特に生物進化において視覚・聴覚より早期に発達した感覚のため、他の感覚と比較しても相対的に香りを嗅いだ際に「それが何か？」を認識する前に、好き嫌いや快不快の刺激がダイレクトに脳辺縁系へ伝えられ、身体調節に至るという特殊性を有すると考えられている。また、使用した香りが個人にとっての好き嫌い、嗅いだ経験の有無によって中枢への刺激部位が変わることも知られている。嫌いな香りや慣れ親しまない香りには生体防衛的な反応が生じるからである。

特に、本研究は対象が全員高齢の日本人で、必ずしも使用された香りは日本人に馴染みのある一般的な香りとは言えなかったため、期待された効果が得られなかった可能性も否定しきれない。

しかし、好みや快不快を考慮したランダム化比較試験は現実には壁に突き当たるのが実情である。好き嫌いという個別性を優先してしまうと無作為でなくなってしまう上、好みのタイプ別に分けたとしても統計的検出力の関係で標本数が膨れ上がってしまうため、現実的に実行可能な規模での検証試験では難しくなってしまうというダブルスタンダードが存在するためである。

この議論はアロマ業界でも以前から存在していたが、筆者らは認知機能に関しては過去の先行研究が、一律に全員に同じ内容の香りが提供されて結果を得ていたこともあり、本研究でも同様の効果を期待して研究がデザインされた。

認知機能や心理尺度の結果に対して、事後分析の結果から、さらににおいの好き嫌いを共変量として主要評価項目である PASAT の結果を調整しても結果は変わらなかった。しかし今後は、本人の嗜好や体質を考慮したいいわゆるテ일러メイドなアプローチがより望ましいということの検証は必要と考えられる。

現実には自らアロマセラピーを積極的に嗜む者や施術者の調合下で本格的にアロマセラピーを積極的に受ける者は当然、個人の香りの嗜好性や体質を考慮し精油を選択する。一方で、「認知機能の悪化予防や改善が期待されるから」との理由で行われるアロマセラピーの多くは、得てして当事者は殆どアロマなどに関心がないことも珍しくなく、当事者の嗜好性や体質よりは家族の期待や意向が優先される形で行われことが少なくない。また介護施設などでのアクティビティーとしてのアロマセラピーでも、一人ひとりに合わせて精油を調合されるというよりは、スタッフの好みが優先されたり、「不眠にいいだろうから」「認知症に効くはずだから」と対症療法的に一律の香りを参加者に結果的に強いられることも少なくはないのでないだろう。

本研究だけでは、確定的結論を下すことは到底できないが、一律の嗅覚刺激を当然の前提として、「科学的だから」と一部だけの側面を見て猫も杓子も「集団を対象とするランダム化比較試験」という現状の学術的潮流は、逆に倫理的問題にすら抵触するのではないかと考えられた。つまり、自戒を込めて言うならば認知機能の低下者に対しては自発的なアロマセラピーだけではなく、周囲の期待から当事者の個別性を

十分に考慮されないままで多くのアロマセラピーの実践がなされているのが「認知症アロマ研究」の実情だからである。

これは漢方薬での議論にも共通する。近年、経験医学である漢方薬が現行の西洋医学に適応される集団対象でのランダム化比較試験で期待されたほどの結果が出ず、漢方薬が過小評価されがちとの懸念は多い。その理由として、症状に対して科学的アルゴリズムという名目で単純一律な投与がなされる傾向が強い対症療法の西洋薬とは異なり、同じ頭痛であっても証などの個人の体質や体力、症状などの個別性をしっかりと見極めた上での処方でなければ、かえって有害無益であるというのが漢方医学やさらにその源流に位置する中医学の治療観だからである。

個の適性に対して、場合分けを考慮した集団対象のランダム化比較試験は必然的に大規模にならざるを得ず、コストが莫大になり非現実的である。結果、本来の治療観を無視した漢方医学では控えるべきであるはずの対症療法に準じてデザインされたランダム化比較試験が実際的には少なくないという潮流は、決して看過できない。例えば、認知症の不穏状態だからと証など全く考慮されずに、安易に抑肝散が一律投与されるケースも医療現場では少なくないと聞く。しかしその背後で、体全体のバランスの崩れに気づかぬまま、潜在的有害事象が生じている危険性は強く憂慮される。

アロマセラピーも、一律提供されるというのは本来の趣旨とは異なるものである。特に、楽しみや癒やしという側面も大きいアロマセラピーの性質を考えても、利用者の好みや快不快（それ以外の多種多様な要因もだが）こそが優先されるべきとされており、それが効果に寄与するとの報告も認める。しかし、多くの集団対象でのランダム化比較試験では一律投与によって効果検証されていることは、結果的にアロマセラピーの効能を過小評価し、有害事象の過大評価につながりかねないと筆者は考える。それは芸術療法や音楽療法が、画一的に提供されて受療者が不快を感じることなどは通常は起こり得ず、セラピストによって個別性に考慮し、適応する香りを見極められてこそ効果が最大化されるという点では、他の非医学的介入の効果検証においても共通して見つめ直されるべき点かもしれない。

不適切なランダム化比較試験による介入の過小評価は問題である一方で、医学だけでなく健康食品等においてまで集団対象でのランダム化比較試験が重視され求められているのも時代の流れではある。効果を謳う以上、提供者にはその科学的根拠が強く求められるのもまた事実である。どのようにこれらに折り合いをつけていくかの議論は以前から存在してきたものではあるが、本研究を通じて、集団レベルの介入効果としての平均処置効果（ATE：average treatment effect）のみが重視され、個人における治療効果（ITEs：individual treatment effects）や治療効果の不均一性（heterogeneity of treatment effects）を無視した集団対象でのランダム化比較試験一辺倒になりがちな昨今の潮流に対して、改めてその整合性をどう見つけるのかという課題を再認識したというのが本研究を通じての感想ではある。筆者らは、その点について、患者の数値化しにくい特質にもとづく経験医学である漢方医学界が目指す現在のスタンスには学ぶところが多いと考える。

そして、そこにはエビデンスとしては低いとみなされがちな“症例報告（case report）”をより良質なものにして集積していくという方向性や、ランダム化比較試験

ではあるが 1 人の対象者に対してランダムにいろいろな介入を与えて、それぞれの介入効果を統計学的に評価する研究デザインである“N of 1 試験”による非集団対象でのランダム化比較試験とその集積などが期待されると考えられる。特に、症例報告は「因果」に関しては最も弱いエビデンスとされ、それだけでは全体評価での信頼性は高いとはみなされない。しかし、「何が起きたか」については信頼性が高いことも事実で、精緻な観察によって得られた知見の重要性は決して過小評価されるものではない[52]。症例報告ではあっても、具体的にイメージが持ちやすく、高い再現性をもって応用可能という側面は特に実践の上では、結果を出すのに都合の良い患者だけを集めることが許される効能研究の結果だけが集積されて治療指針が作成され、実臨床とは乖離した **evidence-practice gap** に陥るリスクを生じやすい集団対象でのランダム化比較試験に、相補的存在となりうることも指摘されている。世界的に先進的な疫学教育の領域では、決してランダム化比較試験のような量的研究だけに偏るのではなく、学術的方法論に基づいた質的研究の重要性も強調されているが、質の高い症例報告もその代表的なものである。こうした質の高い一つ一つの症例報告の集積こそが、結果的には AI 時代を迎える中での質の高いビッグデータ時代での成功に直結する視点とも言える。アロマセラピーにおいても、こういった流れを踏まえた上での建設的な効果検証が今後も重ねられることを期待する。

限界：

本研究の限界は、いくつか挙げられる。一つ目には、参加募集対象者がリサーチ会社登録者であることから、今回得られた結果の一般化に際しては十分な注意が必要と考えられた。また、本文中で指摘しているように対象者が健常高齢者であったために、天井効果で本来期待された効果が得られなかった可能性が生じたことで、アロマセラピーの過小評価に至る危険性には注意が必要と考えられる。また、3 つ目としてはこれも本文中で何度も指摘したように、探索的位置づけの研究ということから多くの副次的評価項目を設定したことから、統計的多重性の問題が懸念される点である。つまり、副次的評価項目で有意だったウェルビーイングに効果がなく、嗅覚を悪化させるという結果は偽陽性の危険性が捨てきれない。一方で、主要評価項目の結果についてはボンフェローニ補正後も有意差は堅持されることから、PASAT-2 秒条件のアロマセラピーによる効果について堅固なものであったと言える。

5. 結論

ランダム化比較試験の結果、健常高齢者に 3 か月間にわたりアロマセラピーを行った結果、注意機能を評価する PASAT-2 秒条件においてプラセボ群に対して有意な改善を認めた。一方で、他の認知機能や心理尺度では差は認めなかった上、ウェルビーイングと嗅覚同定能では、プラセボ群が優れていた。注意機能は、高齢者の自立生活の維持に記憶や他の認知機能以上に重要と考えられてきておりその臨床的意義は大きい。一方で、今回効果が示されなかった記憶をはじめとする他の認知機能や心理社会的側面に効果が得られるためのアロマセラピー自体の工夫、嗅覚低下を生じない適切な強

度等の同定、そしてより望ましいアロマセラピーの疫学的方法論の検討等、課題は多く残されており、今後のさらなる検証が望まれる。

6. 利害相反

本演題に関して、筆頭研究者に開示すべき利益相反はない。

7. 謝辞

本研究の遂行にあたり、多岐にわたるご助言を頂いた社会福祉法人千種会 Les、株式会社再生抗加齢科学研究所、および慶應義塾大学医学部精神神経科学教室のみなさま方に深く感謝申し上げます。

8. 引用文献

1. Shikimoto, R., M. Sado, and M. Mimura, [*The Social Costs of Dementia in Japan: Focusing on the Informal Care Cost*]. Brain Nerve, 2016. **68**(8): p. 939-44.
2. Grant, C., *Heard on the Street :Alzheimer's Setback Is Warning to Biogen Investors in The Wall Street Journal*. 2017, Dow Jones & Company, Inc. : New York.
3. Devanand, D.P., *Olfactory Identification Deficits, Cognitive Decline, and Dementia in Older Adults*. Am J Geriatr Psychiatry, 2016. **24**(12): p. 1151-1157.
4. Braak, H. and E. Braak, *Neuropathological staging of Alzheimer-related changes*. Acta Neuropathol, 1991. **82**(4): p. 239-59.
5. Gold, G., et al., *Clinical validity of Braak neuropathological staging in the oldest-old*. Acta Neuropathol, 2000. **99**(5): p. 579-82.
6. Devanand, D.P., et al., *Olfactory deficits predict cognitive decline and Alzheimer dementia in an urban community*. Neurology, 2015. **84**(2): p. 182-9.
7. Jimbo, D., et al., *Specific feature of olfactory dysfunction with Alzheimer's disease inspected by the Odor Stick Identification Test*. Psychogeriatrics, 2011. **11**(4): p. 196-204.
8. Doty, R.L., *Olfaction in Parkinson's disease and related disorders*. Neurobiol Dis, 2012. **46**(3): p. 527-52.
9. Macklis, J.D., *Neurobiology: New memories from new neurons*: Nature. 2001 Mar 15;410(6826):314-5, 317. doi: 10.1038/35066661.
10. Shors, T.J., et al., *Neurogenesis in the adult is involved in the formation of trace memories*. Nature, 2001. **410**(6826): p. 372-6.
11. Peters, J.M., et al., *Olfactory function in mild cognitive impairment and Alzheimer's disease: an investigation using psychophysical and electrophysiological techniques*. Am J Psychiatry, 2003. **160**(11): p. 1995-2002.
12. Forrester, L.T., et al., *Aromatherapy for dementia*. Cochrane Database Syst Rev, 2014. **25**(2).
13. Jimbo, D., et al., *Effect of aromatherapy on patients with Alzheimer's disease*. Psychogeriatrics, 2009. **9**(4): p. 173-9.
14. Kouzuki, M., et al., *Evaluation of the effect of aroma oil as a bath salt on cognitive function*. Psychogeriatrics, 2020. **20**(2): p. 163-171.

15. Snow, L.A., L. Hovanec, and J. Brandt, *A controlled trial of aromatherapy for agitation in nursing home patients with dementia*. J Altern Complement Med, 2004. **10**(3): p. 431-7.
16. Moss, M., et al., *Modulation of cognitive performance and mood by aromas of peppermint and ylang-ylang*. Int J Neurosci, 2008. **118**(1): p. 59-77.
17. Moss, M., et al., *Aromas of rosemary and lavender essential oils differentially affect cognition and mood in healthy adults*. Int J Neurosci, 2003. **113**(1): p. 15-38.
18. 川畑葉月 and 上出直人, *ラベンダー精油によるアロマセラピーは健常成人の認知機能を高める*. Aroma Research, 2014. **15**(2): p. 174-178.
19. 和氣, 大., et al., *アロマセラピーは健常高齢者の認知機能を改善するか*. 老年精神医学雑誌, 2016. **27**(増刊 II): p. 166.
20. Birte-Antina, W., et al., *Olfactory training with older people*. Int J Geriatr Psychiatry, 2018. **33**(1): p. 212-220.
21. Petersen, R.C., et al., *Memory function in normal aging*. Neurology, 1992. **42**(2): p. 396-401.
22. Petersen, R.C., et al., *Mild Cognitive Impairment: Clinical Characterization and Outcome*. Archives of Neurology, 1999. **56**(3): p. 303-308.
23. Estévez-González, A., et al., *Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging*. Int J Geriatr Psychiatry, 2003. **18**(11): p. 1021-8.
24. Eckerström, C., et al., *A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia*. J Alzheimers Dis, 2013. **36**(3): p. 421-31.
25. Smith, D.G., et al., *Impact of multiple comparisons in randomized clinical trials*. Am J Med, 1987. **83**(3): p. 545-50.
26. Ball, E.L., et al., *Aromatherapy for dementia*. Cochrane Database Syst Rev, 2020. **8**(8).
27. Chamine, I. and B.S. Oken, *Aroma Effects on Physiologic and Cognitive Function Following Acute Stress: A Mechanism Investigation*. J Altern Complement Med, 2016. **22**(9): p. 713-21.
28. Chamine, I. and B.S. Oken, *Aroma Effects on Physiologic and Cognitive Function Following Acute Stress: A Mechanism Investigation*. Journal of alternative and complementary medicine (New York, N.Y.), 2016. **22**(9): p. 713-721.
29. Sasaki, T., et al., *Dentate network activity is necessary for spatial working memory by supporting CA3 sharp-wave ripple generation and prospective firing of CA3 neurons*. Nature Neuroscience, 2018. **21**(2): p. 258-269.
30. Tisserand, D.J. and J. Jolles, *On the involvement of prefrontal networks in cognitive ageing*. Cortex, 2003. **39**(4-5): p. 1107-28.
31. 田淵, 肇., *【高齢者における前頭葉機能障害】 高齢者における遂行機能障害*. 老年精神医学雑誌, 2004. **15**(6): p. 731-736.
32. 目黒謙一, *高齢者にみられる認知障害の特徴】 初期/軽度の認知障害 軽度認知障害(MCI)の問題を中心に*. 老年精神医学雑誌, 2006. **17**(4): p. 379-384.
33. 山鳥重, ed. *高次脳機能障害といかに生きるかー神経心理学の立場から*. 2 版 ed. 高次脳機能障害者の世界, ed. 山田規敏子. 2013, 協同医書出版.
34. Ames, D., A. Burns, and J. O'Brien, *Dementia 3Ed*. 2005: Taylor & Francis.
35. Sohlberg, M.K.M. and C.A. Mateer, *Introduction to Cognitive Rehabilitation: Theory and remediation of attention disorders*. Sohlberg, M.K.M.

Mateer, C.A. ed. 1989, New York: Guilford Publications.

36. 船山, 道., 【認知症の診断・臨床・治療とケア、認知症の人の *Well-being*】 症候学 認知症の認知機能障害を考える 認知症と遂行機能障害. 老年精神医学雑誌, 2016. **27**(増刊 I): p. 53-60.
37. Dawson, J.D., et al., *Neuropsychological predictors of driving errors in older adults*. Journal of the American Geriatrics Society, 2010. **58**(6): p. 1090-1096.
38. Hedges, L.V., I. Olkin, and J.A. Olkin, *Statistical Methods for Meta-Analysis*. 1985: Elsevier Science.
39. Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*. 2013: Taylor & Francis.
40. Thompson, B., *Effect sizes, confidence intervals, and confidence intervals for effect sizes*. Psychology in the Schools, 2007. **44**(5): p. 423-432.
41. Albert, M.S., et al., *The diagnosis of mild cognitive impairment 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**(3): p. 270-9.
42. Hodges JR, 臨床家のための高次脳機能のみかた. 2011: 新興医学出版社.
43. Lezak MD, レザック神経心理学的検査集成. 2005: 創造出版.
44. Sánchez-Vidaña, D.I., et al., *The Effectiveness of Aromatherapy for Depressive Symptoms: A Systematic Review*. Evid Based Complement Alternat Med, 2017. **5869315**(10): p. 4.
45. Coelho, H.F., K. Boddy, and E. Ernst, *Massage therapy for the treatment of depression: a systematic review*. Int J Clin Pract, 2008. **62**(2): p. 325-33.
46. 宗未来, et al., 精神科領域におけるアロマセラピー:精神科領域における統合医療と治療における新しい試み

精神科, 2020. **37**(1): p. 22-28.

47. Lin, P.C., et al., *Effects of aromatherapy on sleep quality: A systematic review and meta-analysis*. Complement Ther Med, 2019. **45**: p. 156-166.
48. Damm, M., et al., *Olfactory training is helpful in postinfectious olfactory loss: a randomized, controlled, multicenter study*. Laryngoscope, 2014. **124**(4): p. 826-31.
49. Konstantinidis, I., et al., *Use of olfactory training in post-traumatic and postinfectious olfactory dysfunction*. Laryngoscope, 2013. **123**(12): p. 4.
50. Sohrabi, H.R., et al., *Olfactory discrimination predicts cognitive decline among community-dwelling older adults*. Transl Psychiatry, 2012. **2**(5): p. 43.
51. Hummel, T., et al., *Effects of olfactory training in patients with olfactory loss*. Laryngoscope, 2009. **119**(3): p. 496-9.
52. 小川真生 and 川村 孝, フォーラム 漢方のエビデンスを『つたえる』EBM 時代の漢方症例報告, in 第60回日本東洋医学会学術総会

E.委. 津谷喜一郎), Editor. 2009, 日本東洋医学会: 東京.

表3 参加者の属性

	アロマセラピー群 (n=60)		エタノール群 (n=59)	
	人数	%	人数	%
性別				
男性	32	53.3	27	45.8
女性	28	46.6	32	54.2
平均年齢（標準偏差）	69.5±3.71		69.5±3.67	
平均[MMSE]点数（標準偏差）	28.5±1.59		28.6±1.25	
においスティック検査（点数）				
6	3	5.0	5	8.3
7	7	11.7	4	6.7
8	3	5.0	7	11.7
9	16	26.7	6	10.0
10	11	18.3	13	21.7
11	15	25.0	14	23.3
12	5	8.3	11	18.3
就労状況				
何かしらの仕事をしている	23	38.3	21	35.0
リタイア	30	50.0	26	43.3
専業主婦	14	23.3	19	31.7
最終学歴				
中学	2	3.3	1	1.7
高校	13	21.7	15	25.0
短大・高専	13	21.7	12	20.0
大学・大学院	32	53.3	32	53.3
婚姻状況				
既婚	49	81.7	48	80.0
離婚	4	6.7	8	13.3
死別	3	5.0	2	3.3
未婚・結婚歴なし	5	8.3	2	3.3

現病歴				
著しい嗅覚低下きたす病気はない	60	100	59	100
精神科治療中ではない	59	98.3	59	100.0
治療歴/ノアルはない身体疾患はない	18	30.5	25	41.7
既往歴				
高血圧	16	27.1	12	20.0
脂質異常症	6	10.2	3	5.0
心筋梗塞	0	0.0	0	0.0
狭心症	1	1.7	1	1.7
不整脈	1	1.7	2	3.3
脳梗塞	0	0.0	0	0.0
脳出血	0	0.0	0	0.0
糖尿病	5	8.5	2	3.3
痛風/高尿酸血症	1	1.7	1	1.8
睡眠時無呼吸症候群	1	1.7	0	0.0
喘息	0	0.0	1	1.7
パーキンソン病	0	0.0	0	0.0
うつ病	0	0.0	0	0.0
PTSD	0	0.0	0	0.0
統合失調症	0	0.0	0	0.0
双極性障害	0	0.0	0	0.0
その他	33	55.9	27	45.0
世帯年収				
100万円以下	2	3.3	0	0.0
100～500万円	39	65.0	35	58.3
500～1000万円	17	28.3	20	33.3
1000万円以上	2	3.3	5	8.3
飲酒歴				
飲まない	22	36.7	21	35.0
毎日	13	21.7	9	15.0
5～6日/週	4	6.7	4	6.7
週の半分	7	11.7	8	13.3
1～2日/週	8	13.3	9	15.0
数日/月	6	10.0	9	15.0
アルコールや他の薬物依存	0	0.0	0	0.0
喫煙歴				
今も吸う	0	0.0	0	0.0
習慣なし	29	48.3	45	76.3
かつて吸った	31	51.7	14	23.7
喫煙本数（本）	19.4±10.99		17.3±9.36	
喫煙期間（年）	19.6±13.89		18.5±10.40	
睡眠習慣				
睡眠時間	6.7±0.96		6.6±0.92	

表 4 介入前の評価指標

	ベースライン					
	アロマ群		プラセボ群		全体	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
PASAT_2秒条件_正答数	32.97	14.22	35.03	14.11	34.01	14.15
PASAT_1秒条件_正答数	21.53	8.16	20.78	8.37	21.15	8.24
SDMT_点数	52.20	8.32	52.43	7.33	52.32	7.77
RAVLT_リストA_想起数合計	46.63	8.76	49.32	10.41	47.98	9.67
RAVLT_リストA_遅延再生（干渉後）_正答数	9.90	2.83	9.97	3.30	9.93	3.07
RAVLT_リストA_再認_正答数	14.33	1.14	14.28	1.09	14.31	1.11
MMSE点	28.48	1.59	28.58	1.25	28.53	1.43
においスティックテスト_正解数	9.47	1.64	9.73	1.86	9.60	1.75
PHQ-9	1.82	2.34	1.45	2.21	1.63	2.27
GAD-7	1.81	2.72	1.33	2.32	1.57	2.52
SDS_仕事学業	1.43	0.81	1.48	1.19	1.46	1.01
SDS_社会生活	1.40	1.03	1.40	1.14	1.40	1.08
SDS_家庭内のコミュニケーションや役割	1.37	1.10	1.47	1.16	1.42	1.13
ウェルビーイング（CES-D4項目）	8.05	3.40	7.97	3.15	8.01	3.26
一般的信頼尺度	31.00	6.60	31.85	5.13	31.43	5.90
睡眠_PSQI質問票総合得点	4.10	2.10	3.85	2.24	3.98	2.17

表 5 介入効果の検討：におい刺激による神経心理学的検査項目の時系列変化に与える影響（線形混合モデル）

	介入群		プラセボ群		群間差(介入群-プラセボ群)		
	LS mean	95%CI	LS mean	95%CI	LS mean	95%CI	P-value
PASAT_2秒条件_正答数							
Time 1	32.94	29.32 , 36.56	35.41	31.79 , 39.02			
Time 2	38.74	35.11 , 42.37	37.88	34.27 , 41.49			
△Time2-Time1	5.80	3.76 , 7.84	2.48	0.48 , 4.47	3.32	0.47 , 6.17	0.023
PASAT_1秒条件_正答数							
Time 1	21.52	19.25 , 23.79	20.88	18.62 , 23.14			
Time 2	22.58	20.30 , 24.86	23.12	20.86 , 25.38			
△Time2-Time1	1.06	-0.81 , 2.93	2.24	0.41 , 4.07	-1.18	-3.80 , 1.44	0.374
RAVLT_想起数合計							
Time 1	46.63	44.01 , 49.24	49.19	46.57 , 51.80			
Time 2	48.27	45.66 , 50.89	53.14	50.52 , 55.75			
△Time2-Time1	1.64	-0.80 , 4.09	3.95	1.51 , 6.39	-2.31	-5.76 , 1.15	0.189
RAVLT_想起数(第5試行)							
Time 1	11.56	11.00 , 12.12	11.88	11.32 , 12.44			
Time 2	11.88	11.32 , 12.45	12.25	11.70 , 12.81			
△Time2-Time1	0.32	-0.15 , 0.80	0.37	-0.09 , 0.84	-0.05	-0.71 , 0.61	0.881
RAVLT_想起数(第5試行-第1試行)							
Time 1	6.00	5.46 , 6.54	5.73	5.19 , 6.26			
Time 2	4.92	4.38 , 5.45	5.05	4.52 , 5.59			
△Time2-Time1	-1.09	-1.75 , -0.42	-0.68	-1.34 , -0.01	-0.41	-1.35 , 0.53	0.394
RAVLT_遅延再生(干渉後)							
Time 1	9.97	9.22 , 10.71	9.95	9.21 , 10.69			
Time 2	10.52	9.77 , 11.27	10.62	9.88 , 11.37			
△Time2-Time1	0.56	-0.12 , 1.23	0.67	0.00 , 1.34	-0.12	-1.07 , 0.84	0.811
RAVLT_再認							
Time 1	14.22	13.89 , 14.56	14.19	13.86 , 14.52			
Time 2	13.93	13.59 , 14.26	14.34	14.01 , 14.67			
△Time2-Time1	-0.30	-0.62 , 0.03	0.15	-0.17 , 0.47	-0.45	-0.91 , 0.01	0.054
SDMT_点数							
Time 1	52.03	49.96 , 54.10	52.60	50.57 , 54.63			
Time 2	52.89	50.86 , 54.92	52.98	50.97 , 55.00			
△Time2-Time1	0.86	-0.42 , 2.15	0.39	-0.80 , 1.57	0.48	-1.27 , 2.23	0.589
MMSE							
Time1	28.48	28.12 , 28.85	28.58	28.22 , 28.95			
Time2	28.41	27.89 , 28.93	28.86	28.34 , 29.39			
△Time2-Time1	-0.08	-0.63 , 0.48	0.28	-0.28 , 0.84	-0.35	-1.14 , 0.43	0.375
においスティックテスト(正解数)							
Time 1	9.50	9.00 , 10.00	9.73	9.23 , 10.24			
Time 2	8.96	8.45 , 9.48	9.95	9.45 , 10.45			
△Time2-Time1	-0.54	-0.95 , -0.12	0.22	-0.19 , 0.62	-0.75	-1.34 , -0.17	0.012

LS mean: Least squares means; 95%CI: 95% confidence interval; △: 変化量。

解析方法: 各評価項目を従属変数とする線形混合モデルにて解析。

固定効果に、群(介入群 or プラセボ群)、測定時期(Time1 or Time2)、及びその交互作用項(群*測定時期)、
変量効果に、Subjects、反復効果に測定時期を設定した。

表6 介入効果の検討：におい刺激による心理学的検査項目の時系列変化に与える影響（線形混合モデル）

	介入群		プラセボ群		群間差(介入群-プラセボ群)		
	LS mean	95%CI	LS mean	95%CI	LS mean	95%CI	P-value
PHQ-9							
Time 1	1.82	1.17, 2.46	1.45	0.80, 2.10			
Time 2	2.13	1.48, 2.79	1.82	1.17, 2.46			
△Time2-Time1	0.32	-0.29, 0.92	0.37	-0.23, 0.97	-0.05	-0.91, 0.80	0.905
GAD-7							
Time 1	1.93	1.27, 2.60	1.42	0.75, 2.08			
Time 2	1.72	1.05, 2.38	1.27	0.60, 1.93			
△Time2-Time1	-0.22	-0.71, 0.27	-0.15	-0.64, 0.34	-0.07	-0.76, 0.63	0.850
SDISS_仕事/学業							
Time 1	1.43	1.15, 1.71	1.48	1.20, 1.76			
Time 2	1.49	1.21, 1.77	1.63	1.35, 1.91			
△Time2-Time1	0.06	-0.14, 0.25	0.15	-0.04, 0.34	-0.09	-0.37, 0.18	0.497
SDISS_社会生活							
Time 1	1.40	1.11, 1.69	1.40	1.11, 1.69			
Time 2	1.35	1.06, 1.64	1.52	1.23, 1.80			
△Time2-Time1	-0.05	-0.30, 0.21	0.12	-0.13, 0.37	-0.16	-0.52, 0.19	0.366
SDISS_家庭内のコミュニケーションや役割							
Time 1	1.37	1.08, 1.65	1.47	1.18, 1.75			
Time 2	1.27	0.98, 1.56	1.60	1.31, 1.89			
△Time2-Time1	-0.10	-0.37, 0.17	0.13	-0.13, 0.40	-0.23	-0.61, 0.14	0.226
ウェルビーイング(CES-D4項目)							
Time 1	12.05	11.27, 12.83	11.97	11.19, 12.74			
Time 2	11.86	11.07, 12.65	12.90	12.12, 13.68			
△Time2-Time1	-0.19	-0.95, 0.56	0.93	0.19, 1.68	-1.13	-2.19, -0.06	0.038
一般信頼尺度							
Time 1	31.00	29.46, 32.54	31.85	30.31, 33.39			
Time 2	31.28	29.73, 32.84	32.38	30.84, 33.93			
△Time2-Time1	0.28	-0.93, 1.50	0.53	-0.67, 1.73	-0.25	-1.96, 1.46	0.772
睡眠_睡眠の質							
Time 1	1.07	0.92, 1.22	0.95	0.80, 1.10			
Time 2	1.01	0.85, 1.16	0.90	0.75, 1.05			
△Time2-Time1	-0.06	-0.20, 0.08	-0.05	-0.19, 0.09	-0.01	-0.21, 0.19	0.917
睡眠_入眠時間							
Time 1	0.63	0.44, 0.83	0.45	0.25, 0.65			
Time 2	0.55	0.36, 0.75	0.40	0.20, 0.60			
△Time2-Time1	-0.08	-0.23, 0.08	-0.05	-0.20, 0.10	-0.03	-0.25, 0.19	0.788
睡眠_睡眠時間							
Time 1	1.14	0.96, 1.33	1.15	0.96, 1.34			
Time 2	1.16	0.97, 1.35	1.18	1.00, 1.37			
△Time2-Time1	0.01	-0.15, 0.17	0.03	-0.12, 0.19	-0.02	-0.25, 0.21	0.860
睡眠_睡眠効率							
Time 1	0.17	0.07, 0.28	0.12	0.02, 0.23			
Time 2	0.12	0.01, 0.23	0.20	0.09, 0.31			
△Time2-Time1	-0.06	-0.18, 0.07	0.08	-0.05, 0.20	-0.13	-0.31, 0.05	0.150
睡眠_睡眠困難							
Time 1	0.77	0.63, 0.90	0.72	0.58, 0.85			
Time 2	0.82	0.69, 0.96	0.73	0.60, 0.87			
△Time2-Time1	0.06	-0.10, 0.21	0.02	-0.14, 0.17	0.04	-0.18, 0.26	0.711
睡眠_眠剤の使用							
Time 1	0.10	-0.03, 0.23	0.12	-0.01, 0.24			
Time 2	0.12	-0.01, 0.25	0.10	-0.03, 0.23			
△Time2-Time1	0.02	-0.08, 0.12	-0.02	-0.12, 0.08	0.03	-0.11, 0.18	0.629
睡眠_日中覚醒困難							
Time 1	0.18	0.05, 0.32	0.35	0.21, 0.49			
Time 2	0.32	0.19, 0.46	0.35	0.21, 0.49			
△Time2-Time1	0.14	0.00, 0.29	0.00	-0.14, 0.14	0.14	-0.06, 0.35	0.174
睡眠_総合得点							
Time 1	4.10	3.51, 4.70	3.85	3.26, 4.44			
Time 2	4.09	3.50, 4.69	3.87	3.28, 4.46			
△Time2-Time1	-0.01	-0.48, 0.46	0.02	-0.45, 0.48	-0.03	-0.69, 0.63	0.932

LS mean: Least squares means; 95%CI: 95% confidence interval; △: 変化量。

解析方法: 各評価項目を従属変数とする線形混合モデルにて解析。

固定効果に、群(介入群 or プラセボ群)、測定時期(Time1 or Time2)、及びその交互作用項(群*測定時期)、

変量効果に、Subjects、反復効果に測定時期を設定した。