Validation, determinants and impact of medication use

Dissertation

zur Erlangung des Doktorgrades (PhD)

der Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

Nersi Alaeddin

aus Teheran, Iran

2024

Angefertigt mit der Genehmigung

der Medizinischen Fakultät der Universität Bonn

1. Gutachterin: Prof. Dr. Dr. Monique M. B. Breteler
2. Gutachterin: Prof. Dr. med. Julia C. Stingl
3. Mitglied: Prof. Dr. Dr. Robin P. Peeters
4. Mitglied: Prof. Dr. Andreas Mayr

Tag der mündlichen Prüfung: 20. März 2024

Aus dem Deutschen Zentrum für Neurodegenerative Erkrankungen (DZNE e.V.), Populationsbezogene Gesundheitsforschung

Direktorin: Prof Dr. Dr. Monique M.B. Breteler

Table of content

List of abbreviations

ACB-Score	Anticholinergic Burden Score
ADR	Adverse drug reaction
ANOVA	Analysis of variance
APP	Amyloid-β precursor protein
ATC	Anatomical Therapeutic Code
BMI	Body Mass Index
CI	Confidence interval
CKD	Chronic kidney disease
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
DTI	Diffusion tension imaging
DZNE	German Center for Neurodegenerative Diseases
EDTA	Ethylenediaminetetraacetic acid
eGFR	Estimated glomerular filtration rate
EMR	Electronic medical record
ESC	European Society of Cardiology
FA	Fractional anisotropy
HbA1c	Glycated haemoglobin
HCT	Hydrochlorothiazide
HRT	Hormone replacement therapy (including oral contraceptives)
ICH-GCP	International Council for Harmonization Good Clinical Practice
IDOM	Instrument zur datenbankgestützten Online-Erfassung von Medikamenten
IMBIE	Institute for Medical Biometry, Informatics and Epidemiology

ISCED-11	International Standard Classification of Education
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LIMS	Laboratory Information Management System
LT4	Levothyroxine
M	Mean
m/z	Mass-to-charge ratio
MD	Mean diffusivity
MRI	Magnetic Resonance Imaging
N	Number of participants
NSAID	Nonsteroidal anti-inflammatory drugs
OR	Odds ratio
OTC	Over-the-counter medication
PPI	Proton pump inhibitor
PSS	Perceived Stress Scale
Ref.	Reference group
SBP	Systolic blood pressure
SD	Standard deviation
TH	Thyroid hormone
TSH	Thyrotropin
VLMT	Verbal Learning and Memory Test
vs.	Versus
WM	White matter
WHO	World Health Organization

1 Abstract

In recent decades, global medication consumption has increased substantially due to an aging population and more emphasis on preventative treatment. As medication use increases, concerns naturally arise regarding the potential for unintended adverse effects and their broader implications for health. In response to these concerns, this thesis emphasizes the importance of population-based cohorts in pharmacoepidemiology.

Using data from the ongoing, community-based Rhineland Study, this thesis pursued three main objectives: (I) to validate the reliability of self-reported medication data in the Rhineland Study, (II) to explore the prevalence and putative determinants of over- and undertreatment with commonly used drugs, and (III) to examine associations between proton pump inhibitors (PPIs) and brain functioning.

Validation of self-reported medication use, confirmed through measured plasma drug metabolites, revealed high concordance rates independent of age and sex. This attests to the reliability of the collected medication data. Exploration of over- and undertreatment of commonly used drugs, namely levothyroxine (LT4) and antihypertensive drugs, identified a high proportion of individuals with inadequate treatment. Among LT4 users, comprising almost a quarter of the population, 4% were undertreated, and 18% were overtreated, particularly affecting the elderly. For antihypertensive drugs, 20% were overtreated, and 33% were undertreated, revealing notable sex differences, with women more likely to be overtreated than men. Risk factors such as age, body mass index, and chronic kidney disease influenced over- and undertreatment risks differently in women and men. Lastly, a focused investigation of PPI effects on brain function, utilizing a detailed cognitive battery and state-of-the-art MRI, revealed potential cognitive implications and disruptions in white matter integrity, particularly among younger long-term users.

The high prevalence of inadequate treatment emphasizes the need for cautious dosing and deintensification strategies, with a call for vigilance against overtreatment. Furthermore, this thesis challenges the assumed safety of PPIs, highlighting potential detrimental effects on brain functioning. Recognizing the risk of harm from seemingly safe drugs, prudent use and prescribing practices are crucial. Hence, population-based cohorts in pharmacoepidemiology play a vital role in unravelling patterns and refining treatment guidelines.

2 Introduction and aims

In an era of improved access to effective treatments and a growing focus on preventive healthcare (Guthrie *et al.*, 2015; Moßhammer *et al.*, 2016), pharmacoepidemiology has emerged as an important cornerstone of both public health and medical research. This interdisciplinary field bridges pharmacology and epidemiology, focusing on drug use, the relationship between drug exposure and subsequent (un-)favourable effects in populations (Montastruc *et al.*, 2019). As such, it plays a central role in drug safety research and post-marketing drug surveillance.

Medication use not only enhances the intended therapeutic effects but also increases the potential for unintended effects (Fisher and Welch, 1999). Understanding the long-term (un-)intended effects of drugs typically begins after drug authorization and requires extensive data collection. However, clinical trials and secondary data offer limited insights due to constraints like generalizability, short duration, missing information or time lags. This is where population-based cohorts become essential by offering real-world insights into drug use, implications, and safety profiles in the general population.

One such exemplary study is the ongoing Rhineland Study, a population-based cohort comprising adults aged ≥30 years in Bonn, Germany (Breteler and Wolf, 2014). The overarching aim is to investigate the aetiology and pathogenesis of age-related (neurodegenerative) diseases through in-depth phenotyping by collecting comprehensive data on cardiovascular health, brain imaging, cognition, neurological function, and medication use, among others. Therefore, the Rhineland Study provides a unique opportunity to explore drug utilization and identify potential risks in a real-world setting.

2.1 Validation of drug exposure

Pharmacoepidemiologic studies utilize different data sources to assess drug exposure, each with its own strengths and limitations. While secondary data from electronic medical records (EMRs) are often considered as the gold standard, large population-based cohort studies, such as the Rhineland Study, often rely on self-reported medication use. Concerns persist about the accuracy and reliability of self-reported medication data, hinging on participants' veracity, potentially introducing bias (Althubaiti, 2016).

However, the accuracy of EMRs in reflecting actual medication use remains uncertain. EMRs are primarily designed for medical billing purposes and may not adequately capture actual drug intake (Bots, Groenwold and Dekkers, 2022). This is particularly concerning, given that 40% of patients deviate from prescribed regimens (Wilke, Müller and Morisky, 2011) - a nuance evaded by EMRs, but addressable by self-reported medication use. Ensuring data reliability is paramount to draw informed conclusions, as erroneous or biased findings can have profound implications (West, Ritchey and Poole, 2012).

Although EMRs are often used to validate self-reported drug exposure (Hafferty *et al.*, 2018), their accuracy is limited by inherent limitations. Recent analytical advances present opportunities to overcome these challenges. For instance, one viable approach involves the use of untargeted metabolomics to measure drug metabolites in blood, offering a potential solution to enhance accuracy. Validating self-reported drug use with metabolites can be challenging for drugs taken irregularly or with short half-lives, such as analgesics (Dennis *et al.*, 2018). Nevertheless, this approach provides a more unbiased estimate, particularly for drugs with sufficiently long half-lives used regularly in chronic conditions.

2.2 Over- and undertreatment

The surge in medication use can partly be attributed to the increasing focus on disease prevention (Guthrie *et al.*, 2015; Moßhammer *et al.*, 2016). In the Rhineland study, 66% used at least one drug regularly, with 16% reporting polypharmacy, increasing to 37% in individuals aged ≥65 years (de Vries, Stingl and Breteler, 2021). While medication use is essential for treatment, its increase raises questions about the adequacy of drug use.

Effective treatment involves mitigating the risks of over- and undertreatment, which are associated with adverse clinical outcomes and increased morbidity (Kearney, Treadwell and Marshall, 2017). Undertreatment, denoting the failure to prescribe the correct medication or dosage, often results from patient non-adherence, lack of awareness, or physician judgement (Kearney, Treadwell and Marshall, 2017). This can exacerbate disease progression, worsen symptoms, and undermine overall health and quality of life.

The increased focus on prevention can potentially lead to overtreatment in certain subpopulations. While undertreatment has been extensively studied, there is limited exploration of overtreatment. Overtreatment, characterized by excessive medication prescription or intake, is often driven by factors such as overdiagnosis or patient demand

(Huebscher, 1997; Ooi, 2020). This not only inflates drug expenditures, but could also expose patients to unnecessary adverse effects (Kojima *et al.*, 2012; Lyu *et al.*, 2017).

To address these concerns, it is important to investigate the prevalence over- and undertreatment in the general population, especially for commonly used drugs. Worldwide, antihypertensive drug use has risen, with over 60% of individuals with hypertension in Germany receiving treatment (Zhou *et al.*, 2021). Moreover, Germany has almost four times higher thyroid hormone usage (~11%), compared to other European countries (~4%) (Okosieme *et al.*, 2011; Khattak *et al.*, 2016; Kiel *et al.*, 2020; Wouters *et al.*, 2020; Janett-Pellegri *et al.*, 2021). As both over- and undertreatment with these drugs have been associated with negative health outcomes (Lillevang-Johansen *et al.*, 2018; Zhou *et al.*, 2018), it is important not only to assess the prevalence but also to identify individuals at an increased risk.

2.3 Intended and unintended drug effects

Cohort studies are a gateway to investigate the full spectrum of drug effects, including intended therapeutic benefits and unintended consequences. Increased medication use, including potential overtreatment, heightens the risk of adverse drug reactions (ADRs). ADRs are unintended and often harmful reactions to medications, ranging from minor side effects to serious complications (Edwards and Aronson, 2000). These ADRs represent a significant healthcare burden, contributing to approximately 7% of emergency admissions in Germany and the UK, with ADR-related mortality among hospitalised patients ranging from 0.1% to 4.7% (Lavan and Gallagher, 2016; Schurig *et al.*, 2018). Notably, approximately 50% of these ADRs are considered preventable (Lavan and Gallagher, 2016), presenting an appealing target for intervention.

Individuals' susceptibility to ADRs varies based on health status, genetic predisposition, and interactions with other drugs (Kim, Johnson and Derendorf, 2004; Magro, Moretti and Leone, 2012). Although ADRs can be assessed during and after drug authorization, some may only surface after decades of widespread use. A notable example are proton pump inhibitors (PPIs), used for controlling gastric acid-related gastric disorders (Savarino, Di Mario and Scarpignato, 2009; Dharmarajan, 2021). The presumed safety of long-term PPI use has been questioned, with reports suggesting associations between PPIs and cognitive decline, and in some cases, dementia (Akter *et al.*, 2015; Haenisch *et al.*, 2015;

Gomm *et al.*, 2016). Conversely, others have refuted these claims (Wod *et al.*, 2018; Hussain *et al.*, 2020), necessitating further investigation for a conclusive verdict. This issue carries weight given the widespread use of PPIs in both Europe and the USA (Heidelbaugh et al., 2012; Lanas, 2016; Schumock et al., 2016; Johnson et al., 2017). Alarmingly, in 40% of people using PPIs, this might be inappropriate due to prolonged PPI use, drug-drug interactions, and elevated risk-benefit ratios (Heidelbaugh, Goldberg and Inadomi, 2010; Yang and Metz, 2010; Pasina *et al.*, 2011). Recent data also suggest that nearly 2 million people in Germany are living with dementia (Thyrian et al., 2020), highlighting the need to investigate a potential adverse effect of PPIs on dementia risk.

To understand potential effects of PPIs, an exploration into the intricacies of brain structure is needed. The population-based Rhineland Study, with its cognitive test battery designed to assess cognitive performance throughout adulthood (Bönniger, 2021), provides a valuable opportunity to uncover nuanced connections between PPI use and cognitive outcomes. Additionally, state-of-the-art MRI scanning protocols allow an investigation of structural brain changes that may precede changes in cognitive performance. One promising avenue to study microstructural brain parameters is diffusion tensor imaging (DTI), a technique used in the Rhineland Study to assess the diffusion rate of water in brain tissue, revealing subtle variations in tissue integrity. These variations may indicate early structural changes that are indicative of cognitive decline (Luo *et al.*, 2020).

2.4 Thesis outline and aims

Population-based studies are essential for advancing pharmacoepidemiology, providing insights into medication use and health outcomes. This thesis contributes to this field by initially establishing the reliability of self-reported medication data in the population-based Rhineland Study by validating self-reported drug use via measured drug metabolites in blood (**Chapter 3.1**). Subsequently, **Chapters 3.2** and **3.3** focus on investigating the prevalence and putative determinants of over- and undertreatment within two commonly used drug classes (thyroid hormones and antihypertensive drugs) to identify individuals at increased risk. **Chapter 3.4** then explores associations between proton pump inhibitor use and brain functioning. The concluding **Chapter 4** summarizes the findings' significance, underlining the importance of population-based studies in shaping future research directions in pharmacoepidemiology.

2.5 References

Akter, S. *et al.* (2015) 'Cognitive impact after short-term exposure to different proton pump inhibitors: assessment using CANTAB software', *Alzheimer's Research & Therapy*. Alzheimer's Research & Therapy, 7(1), p. 79. doi: 10.1186/s13195-015-0164-8.

Althubaiti, A. (2016) 'Information bias in health research: definition, pitfalls, and adjustment methods', *Journal of Multidisciplinary Healthcare*, 9, p. 211. doi: 10.2147/JMDH.S104807.

Bönniger, M. M. (2021) *Memory and Neurodegeneration Across the Lifespan Meta Miriam Bönniger*. Available at: https://nbn-resolving.org/urn:nbn:de:hbz:5-63979.

Bots, S. H., Groenwold, R. H. H. and Dekkers, O. M. (2022) 'Using electronic health record data for clinical research: a quick guide', *European Journal of Endocrinology*, 186(4), pp. E1–E6. doi: 10.1530/EJE-21-1088.

Breteler, M. M. B. and Wolf, H. (2014) 'The Rhineland Study: A novel platform for epidemiologic research into Alzheimer disease and related disorders', *Alzheimer's & Dementia*, 10(4S_Part_13), p. 520. doi: 10.1016/j.jalz.2014.05.810.

Dennis, K. *et al.* (2018) 'Metabolomics Approach for Validation of Self-Reported Ibuprofen and Acetaminophen Use', *Metabolites*, 8(4), p. 55. doi: 10.3390/metabo8040055.

Dharmarajan, T. S. (2021) 'The Use and Misuse of Proton Pump Inhibitors: An Opportunity for Deprescribing', *Journal of the American Medical Directors Association*. Elsevier Inc., 22(1), pp. 15–22. doi: 10.1016/j.jamda.2020.09.046.

Edwards, I. R. and Aronson, J. K. (2000) 'Adverse drug reactions: definitions, diagnosis, and management.', *Lancet (London, England)*, 356(9237), pp. 1255–9. doi: 10.1016/S0140-6736(00)02799-9.

Fisher, E. S. and Welch, H. G. (1999) 'Avoiding the Unintended Consequences of Growth in Medical Care', *JAMA*, 281(5), p. 446. doi: 10.1001/jama.281.5.446.

Gomm, W. *et al.* (2016) 'Association of Proton Pump Inhibitors With Risk of Dementia', *JAMA Neurology*, 73(4), p. 410. doi: 10.1001/jamaneurol.2015.4791.

Guthrie, B. *et al.* (2015) 'The rising tide of polypharmacy and drug-drug interactions: population database analysis 1995–2010', *BMC Medicine*. ???, 13(1), p. 74. doi: 10.1186/s12916-015-0322-7.

Haenisch, B. *et al.* (2015) 'Risk of dementia in elderly patients with the use of proton pump inhibitors', *European Archives of Psychiatry and Clinical Neuroscience*, 265(5), pp. 419–428. doi: 10.1007/s00406-014-0554-0.

Hafferty, J. D. *et al.* (2018) 'Self-reported medication use validated through record linkage to national prescribing data', *Journal of Clinical Epidemiology*. Elsevier Inc, 94, pp. 132–142. doi: 10.1016/j.jclinepi.2017.10.013.

Heidelbaugh, J. J. *et al.* (2012) 'Overutilization of proton-pump inhibitors: what the clinician needs to know', *Therapeutic Advances in Gastroenterology*, 5(4), pp. 219–232. doi: 10.1177/1756283X12437358.

Heidelbaugh, J. J., Goldberg, K. L. and Inadomi, J. M. (2010) 'Magnitude and economic effect of overuse of antisecretory therapy in the ambulatory care setting.', *The American journal of managed care*, 16(9), pp. e228-34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21250399.

Huebscher, R. (1997) 'Overdrugging and undertreatment in primary health care.', *Nursing outlook*, 45(4), pp. 161–6. doi: 10.1016/s0029-6554(97)90022-1.

Hussain, S. *et al.* (2020) 'No association between proton pump inhibitor use and risk of dementia: Evidence from a meta-analysis', *Journal of Gastroenterology and Hepatology*, 35(1), pp. 19–28. doi: 10.1111/jgh.14789.

Janett-Pellegri, C. *et al.* (2021) 'Prevalence and factors associated with chronic use of levothyroxine: A cohort study', *PLOS ONE*. Edited by A. Lupattelli, 16(12), p. e0261160. doi: 10.1371/journal.pone.0261160.

Johnson, D. A. *et al.* (2017) 'The Safety of Appropriate Use of Over-the-Counter Proton Pump Inhibitors: An Evidence-Based Review and Delphi Consensus', *Drugs*. Springer International Publishing, 77(5), pp. 547–561. doi: 10.1007/s40265-017-0712-6.

Kearney, M., Treadwell, J. and Marshall, M. (2017) 'Overtreatment and undertreatment: time to challenge our thinking', *British Journal of General Practice*, 67(663), pp. 442–443. doi: 10.3399/bjgp17X692657.

Khattak, R. M. *et al.* (2016) 'Monitoring the prevalence of thyroid disorders in the adult population of Northeast Germany', *Population Health Metrics*. Population Health Metrics,

14(1). doi: 10.1186/s12963-016-0111-3.

Kiel, S. *et al.* (2020) 'Frequency of thyroid function tests and examinations in participants of a population-based study.', *BMC health services research*. BMC Health Services Research, 20(1), p. 70. doi: 10.1186/s12913-020-4910-7.

Kim, K., Johnson, J. A. and Derendorf, H. (2004) 'Differences in Drug Pharmacokinetics Between East Asians and Caucasians and the Role of Genetic Polymorphisms', *The Journal of Clinical Pharmacology*, 44(10), pp. 1083–1105. doi: 10.1177/0091270004268128.

Kojima, G. *et al.* (2012) 'Reducing Cost by Reducing Polypharmacy: The Polypharmacy Outcomes Project', *Journal of the American Medical Directors Association*, 13(9), pp. 818.e11-818.e15. doi: https://doi.org/10.1016/j.jamda.2012.07.019.

Lanas, A. (2016) 'We Are Using Too Many PPIs, and We Need to Stop: A European Perspective', *American Journal of Gastroenterology*. Nature Publishing Group, 111(8), pp. 1085–1086. doi: 10.1038/ajg.2016.166.

Lavan, A. H. and Gallagher, P. (2016) 'Predicting risk of adverse drug reactions in older adults', *Therapeutic Advances in Drug Safety*, 7(1), pp. 11–22. doi: 10.1177/2042098615615472.

Lillevang-Johansen, M. *et al.* (2018) 'Over- and Under-Treatment of Hypothyroidism Is Associated with Excess Mortality: A Register-Based Cohort Study', *Thyroid*, 28(5), pp. 566–574. doi: 10.1089/thy.2017.0517.

Luo, C. *et al.* (2020) 'White Matter Microstructural Damage as an Early Sign of Subjective Cognitive Decline', *Frontiers in Aging Neuroscience*, 11(January), pp. 1–11. doi: 10.3389/fnagi.2019.00378.

Lyu, H. *et al.* (2017) 'Overtreatment in the United States', *PLoS ONE*, 12(9), pp. 1–11. doi: 10.1371/journal.pone.0181970.

Magro, L., Moretti, U. and Leone, R. (2012) 'Epidemiology and characteristics of adverse drug reactions caused by drug–drug interactions', *Expert Opinion on Drug Safety*, 11(1), pp. 83–94. doi: 10.1517/14740338.2012.631910.

Montastruc, J.-L. *et al.* (2019) 'What is pharmacoepidemiology? Definition, methods,

interest and clinical applications.', *Therapie*, 74(2), pp. 169–174. doi: 10.1016/j.therap.2018.08.001.

Moßhammer, D. *et al.* (2016) 'Polypharmacy—an Upward Trend with Unpredictable Effects', *Deutsches Ärzteblatt international*, 113(38), pp. 627–633. doi: 10.3238/arztebl.2016.0627.

Okosieme, O. E. *et al.* (2011) 'Adequacy of thyroid hormone replacement in a general population', *Qjm*, 104(5), pp. 395–401. doi: 10.1093/qjmed/hcq222.

Ooi, K. (2020) 'The Pitfalls of Overtreatment: Why More Care is not Necessarily Beneficial', *Asian Bioethics Review*. Asian Bioethics Review, 12(4), pp. 399–417. doi: 10.1007/s41649-020-00145-z.

Pasina, L. *et al.* (2011) 'Prevalence and appropriateness of drug prescriptions for peptic ulcer and gastro-esophageal reflux disease in a cohort of hospitalized elderly', *European Journal of Internal Medicine*. European Federation of Internal Medicine., 22(2), pp. 205–210. doi: 10.1016/j.ejim.2010.11.009.

Savarino, V., Di Mario, F. and Scarpignato, C. (2009) 'Proton pump inhibitors in GORDAn overview of their pharmacology, efficacy and safety', *Pharmacological Research*, 59(3), pp. 135–153. doi: 10.1016/j.phrs.2008.09.016.

Schumock, G. T. *et al.* (2016) 'National trends in prescription drug expenditures and projections for 2016', *American Journal of Health-System Pharmacy*, 73(14), pp. 1058–1075. doi: 10.2146/ajhp160205.

Schurig, A. M. *et al.* (2018) 'Adverse Drug Reactions (ADR) and Emergencies', *Deutsches Ärzteblatt international*, 115(15), pp. 251–258. doi: 10.3238/arztebl.2018.0251.

Thyrian, J. R. *et al.* (2020) 'Die Prävalenz an Demenz erkrankter Menschen in Deutschland – eine bundesweite Analyse auf Kreisebene', *Der Nervenarzt*, 91(11), pp. 1058–1061. doi: 10.1007/s00115-020-00923-y.

de Vries, F. M., Stingl, J. C. and Breteler, M. M. B. (2021) 'Polypharmacy, potentially inappropriate medication and pharmacogenomics drug exposure in the Rhineland Study', *British Journal of Clinical Pharmacology*, 87(7), pp. 2732–2756. doi: 10.1111/bcp.14671.

West, S. L., Ritchey, M. E. and Poole, C. (2012) 'Validity of Pharmacoepidemiologic Drug

and Diagnosis Data', in *Pharmacoepidemiology*. Wiley, pp. 757–794. doi: 10.1002/9781119959946.ch41.

Wilke, T., Müller, S. and Morisky, D. E. (2011) 'Toward Identifying the Causes and Combinations of Causes Increasing the Risks of Nonadherence to Medical Regimens: Combined Results of Two German Self-Report Surveys', *Value in Health*, 14(8), pp. 1092–1100. doi: 10.1016/j.jval.2011.06.002.

Wod, M. *et al.* (2018) 'Lack of Association Between Proton Pump Inhibitor Use and Cognitive Decline', *Clinical Gastroenterology and Hepatology*. The American Gastroenterological Association, 16(5), pp. 681–689. doi: 10.1016/j.cgh.2018.01.034.

Wouters, H. J. C. M. *et al.* (2020) 'Epidemiology of thyroid disorders in the Lifelines Cohort Study (the Netherlands)', *PLOS ONE*. Edited by S. Naitza, 15(11), p. e0242795. doi: 10.1371/journal.pone.0242795.

Yang, Y. and Metz, D. C. (2010) 'Safety of Proton Pump Inhibitor Exposure', *Gastroenterology*. Elsevier Inc., 139(4), pp. 1115–1127. doi: 10.1053/j.gastro.2010.08.023.

Zhou, B. *et al.* (2021) 'Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants', *The Lancet*, 398(10304), pp. 957–980. doi: 10.1016/S0140-6736(21)01330-1.

Zhou, D. *et al.* (2018) 'Uncontrolled hypertension increases risk of all-cause and cardiovascular disease mortality in US adults: the NHANES III Linked Mortality Study', *Scientific Reports*, 8(1), p. 9418. doi: 10.1038/s41598-018-27377-2.

3 Publications

3.1 Validation of self-reported medication use applying untargeted mass spectrometry-based metabolomics techniques in the Rhineland Study

Received: 30 July 2021 | Revised: 27 October 2021 | Accepted: 24 November 2021

DOI: 10.1111/bcp.15175

ORIGINAL ARTICLE

BJCP BRITISH PHARMACOLOGICAL SOCIETY

Validation of self-reported medication use applying untargeted mass spectrometry-based metabolomics techniques in the Rhineland study

Nersi Alaeddin[1] | **Julia C. Stingl**[2] | **Monique M. B. Breteler**[1,3] | **Folgerdiena M. de Vries**[1]

[1]Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

[2]Institute of Clinical Pharmacology, University Hospital of RWTH Aachen, Aachen, Germany

[3]Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Germany

Correspondence
Folgerdiena M. de Vries, PhD, Deutsches Zentrum für Neurodegenerative Erkrankungen e.V., In der Helmholtz-Gemeinschaft, Venusberg-Campus 1, Building 99, 53127, Bonn, Germany.
Email: dianna.devries@dzne.de

Funding information
Federal Institute for Drugs and Medical Devices (BfArM), Grant/Award Numbers: V-15646/68502/2014-2016, V-17746/68502/2018-2020

Aims: To assess the validity of self-reported continuous medication use with drug metabolites measured in plasma by using untargeted mass spectrometric techniques.

Methods: In a population-based cohort in Bonn, Germany, we compared interview-based, self-reported medication intake with drug-specific metabolites measured in plasma (based on participants who completed their study visits between March 2016 and February 2020). Analyses were done stratified by sex and age (<65 years *vs* ≥65 years). Cohen's kappa (κ) statistics with 95% confidence intervals (CI) were calculated.

Results: A total of 13 drugs used to treat hypertension, gout, diabetes, epilepsy and depression were analysed in a sample of 4386 individuals (mean age 55 years, 56.1% women). Eleven drugs showed almost perfect agreement (κ > 0.8), whereas sitagliptin and hydrochlorothiazide showed substantial (κ = 0.8, 95% CI 0.71–0.90) and moderate agreement (κ = 0.61, 95% CI 0.56–0.66), respectively. Frequency of use allowed sex- and age-stratified analyses for eight and nine drugs, respectively. For five drugs, concordance tended to be higher for women than for men. For most drugs, concordance was higher among individuals aged ≥65 years than among individuals aged <65 years, but these age-related differences were not statistically significant.

Conclusion: High concordance rates between self-reported drug use and metabolites measured in plasma suggest that self-reported drug use is reliable and accurate for assessing drug use.

KEYWORDS
mass spectrometry, metabolomics, molecular epidemiology, pharmacoepidemiology, self-reported data, validity

The authors confirm that the Principal Investigator for this paper is Monique M.B. Breteler.

1 | INTRODUCTION

Large population-based cohort studies often depend on self-reported medication use to determine drug exposure. The quality of self-reported data is frequently questioned because it depends on the accuracy and truthfulness of the participant and is prone to several forms of bias.[1–6] Studies showed no difference in the ability to recall medication use by sex, but reported a decrease with age.[7–9] Another method of assessing medication use that is not expected to be influenced by patient characteristics is the use of secondary data from electronic medical records (EMRs). This method is commonly considered a gold standard for determining medication exposure.[1,10–16] However, it can be questioned whether EMRs indeed reflect actual medication use. EMRs are usually based on prescription data from general practitioners or dispensing data from public pharmacy records. Consequently, medication dispensed in hospitals is often incomplete in these records. Additionally, approximately 40% of patients fail to adhere to their medication as prescribed,[17] which is also not reflected in EMRs but can be addressed in self-reported medication use. Furthermore, self-reported medication use has the advantage of including over-the-counter (OTC) medication and dietary supplements.

Several studies validated self-reported medication use with EMRs.[10–16,18–20] The findings of these studies have been conflicting, showing both over- and underreporting, and levels of agreement often varied across medication classes. High concordance rates were reported for cardiovascular drugs ($\kappa > 0.75$),[10,12,16] while results were inconsistent for drugs prone to stigmatization bias, such as psycholeptics or mood stabilizers (concordance rates ranging from 0.52 to 0.75).[10–12,19] It is unknown whether lower concordance in, for example, psychotropic drugs is due to people giving socially desirable answers or that EMRs do not reflect actual intake.

In general, it is questionable whether medication use extracted from EMRs should be favoured over self-reported medication use. Studies comparing multiple drug metabolites measured in blood using liquid chromatography–tandem mass spectrometry techniques (LC–MS/MS) with EMRs reported huge discrepancies between the two.[21,22] In a patient cohort of 821 US adults, prescription records were compared with a drug metabolite panel consisting of 38 different drugs using an LC–MS/MS assay. Only 46% of the drugs assessed were detected and reported. Of the remaining drugs, 23% were detected in blood but not listed in the prescription records, whereas 30% were present in the prescription records but not detected in blood.[21] Another US study reported discrepancies between medical records and measured metabolites in 63% of the patients.[22]

Depending on a sufficiently long half-life in blood, drug metabolites measured in plasma can provide an unbiased estimate of actual drug intake and can be more appropriate than EMRs to validate self-reported medication use. A study using an untargeted metabolomics approach found good concordance between self-reported acetaminophen use and blood plasma metabolites but poor concordance for ibuprofen.[23] Indeed, validating self-reported use with medication metabolites might not be ideal for analgesics, which are often used irregularly on the occurrence of symptoms and have a short half-life. For drugs used regularly for chronic conditions, validation of self-reported medication use with drug metabolites measured in plasma would be more appropriate.

What is already known about this subject

- The quality of self-reported medication data is frequently questioned because it depends on the accuracy and truthfulness of participants.
- Studies validating self-reported medication use with electronic medical records showed conflicting results with both over- and underreporting.
- Validating self-reported medication use with drug metabolites measured in plasma would be more appropriate.

What this study adds

- For most drugs, including antidepressants, the agreement between self-reported chronic medication use and drug metabolites measured in plasma was almost perfect.
- Within our study self-reported drug intake was a reliable and accurate method for assessing medication exposure.

In this study, we aimed to assess the validity of self-reported continuous medication use with drug metabolites measured in plasma with untargeted mass spectrometry-based metabolomics. Furthermore, we assessed whether the validity of self-reported medication use depended on sex or age.

2 | METHODS

2.1 | Study design and setting

We used baseline data from the Rhineland Study, an ongoing prospective population-based cohort study in Bonn, Germany. This single-centre study started recruitment in 2016 and invites all residents (≥30 years) from two geographically defined areas in Bonn. Contact details of eligible participants are provided by the municipality. Participation is possible on invitation only, regardless of current health status. Those unable to sufficiently understand the informed consent are excluded. All participants undergo in-depth phenotyping, including assessment of cardiovascular measures, brain imaging, cognitive testing, neurologic functioning, untargeted metabolomic profiling in plasma and medication use. Approval to undertake the study was granted by the ethics committee of the University of Bonn, Medical Faculty. This study is performed according to the recommendations of the International Council for Harmonisation (ICH) and the Good Clinical Practice (GCP) standards. We obtained written informed consent

from all participants in accordance with the Declaration of Helsinki. Participants were not offered any financial incentives.

2.2 | Medication data collection

Participants were requested to bring the original packages of all drugs (including OTC drugs, excluding homeopathic drugs) and prescribed supplements used currently or used *as needed* during the last year.[24] As-needed OTC drugs taken for <10 days in the preceding year were not registered. Regular drug use was defined as use at a specific dosing interval, eg, daily, every other day, weekly, without regard to symptoms. Drug use on the occurrence of symptoms was classified as as-needed use. Medication data was assessed interview-based, using a software instrument for database-assisted online collection of medication data (Instrument zur Datenbank gestützten Online-Erfassung von Medikamenten, IDOM).[25] The Research Institute of the Federal Association of Regional Statutory Health Insurance Funds in Germany (Wissenschaftliches Institut der Ortskrankenkassen, WIdO) provides a database containing information on all drugs available on the German market, which is linked to the IDOM software.[26] The name, dosage, Anatomical Therapeutic Chemical (ATC) code,[27] type of use (regularly or as needed) and current prescription status were registered for each preparation.

2.3 | Metabolomics measurements

Drug metabolites were measured in plasma using an untargeted analytical approach (described in detail in section 2.4), which allows the identification of up to 44 different drugs from a reference library (Metabolon Inc., Durham, USA). This in-house reference library contains authenticated standards with retention index, mass-to-charge ratio (m/z) and MS/MS spectral data for each metabolite.[28] The panel contains metabolites of commonly used drugs that are representative of a variety of drug classes (Table A1). Blood samples were collected in the morning after a minimum of 10 hours of fasting using a standard operating procedure. Plasma originated from blood collected in ethylenediaminetetraacetic acid (EDTA)-containing vacutainers. Samples were maintained at −80 °C and accessioned into a laboratory information management system (LIMS) to track all results, samples and derived aliquots. All samples were analysed via ultrahigh-performance liquid-phase chromatography and separation coupled with tandem mass spectrometry.[28–30] Raw data were extracted, and the characteristic chromatographic peaks and relative ion concentrations of the detected metabolites were determined for each sample. Subsequently, the spectrometry data were analysed to identify and quantify individual components using the quantify individual components in a sample method.[31] Here, ions for a given metabolite are determined from LC–MS/MS data based on retention time, mass ion intensity and covariance of ion data across the entire sample set. Relative metabolite levels were computed by using the area under the curve. Retention index and mass-to-charge ratio (m/z) for the investigated metabolites are provided in Table A2. Participants were considered users of a specific drug if the respective drug metabolite was detected. Missing values were considered to mirror quantities below levels of detection and were therefore classified as nonusers of the specific drug.

2.4 | Ultrahigh-performance liquid chromatography-tandem mass spectroscopy

We use an LC-MS-based untargeted metabolomics approach for analysing the drug metabolites. All analyses used a Waters ACQUITY ultra-performance liquid chromatograph (UPLC) (Waters Limited, Mississauga, ON, Canada), a Thermo Scientific Q-Exactive high-resolution mass spectrometer interfaced to a heated electrospray ionization source (HESI-II) and an Orbitrap mass analyser with a mass resolution of 35 000 (Thermo Fisher Scientific, Waltham, Massachusetts, USA).[28,32] Sample extracts were dried and reconstituted in solvents compatible with each of the four methods (i-iv) described below. Each reconstitution solvent contained a set of standards (fixed concentrations) to guarantee consistency of injection and chromatography. (i) One aliquot was analysed under acidic positive-ion conditions and chromatographically optimized for more hydrophilic compounds. The extract was gradient eluted from a C18 column (Waters UPLC BEH C18, 2.1 × 100 mm, 1.7 μm) with water and methanol containing 0.05% perfluoropentanoic acid and 0.1% formic acid. (ii) Another aliquot was analysed under acidic positive-ion conditions and chromatographically optimized for hydrophobic compounds. Here, the extract was gradient eluted with methanol, acetonitrile, water, 0.05% perfluoropentanoic acid and 0.01% formic acid from the same C18 column mentioned above and run at a higher overall organic content. (iii) The third aliquot was analysed using basic negative-ion-optimized conditions using a separate C18 column. Basic extracts were gradient eluted from the column with water and methanol with 6.5 mM ammonium bicarbonate (pH 8). (iv) The fourth aliquot was analysed by negative ionization following elution from a HILIC column (Waters UPLC BEH Amide, 2.1 × 150 mm,1.7 μm) using a gradient consisting of water and acetonitrile with 10 mM ammonium formate (pH 10.8).

2.5 | Selection criteria of included drugs

Of all drugs detected with the metabolomics panel, we included only those (i) that were used for long-term conditions and used regularly, (ii) that were self-reported by ≥15 participants and (iii) for which we expected a priori that the available drug metabolites would be appropriate for validation of self-reported drug use based on the pharmacokinetic properties of the drug (assessment by two pharmacists). A total of 75 drug metabolites associated with 44 different drugs (Table A1) can potentially be detected by the metabolomics panel. Of these drugs, we excluded 23 because they are not typically used regularly and six because fewer than 15 participants in our population reported regular use. For the remaining 15 drugs, we assessed

TABLE 1 Selection of drugs based on pharmacokinetic characteristics

Drug class	Drug	t_{max} (h)	Bioavailability (%)	Active metabolite	Measured metabolite(s)	Plasma elimination $t_{½}$ (h) of measured metabolite(s)	Included
Antihypertensive	Enalapril	1	40	Enalaprilat	Enalapril	<1	×
	Hydrochlorothiazide	2–5	70	Excreted unchanged	HCT	6–8	✓
	Metoprolol	1-2	35-50	Metoprolol	(a) Metoprolol (b) Metoprolol acid	3–5	✓
	Candesartan	3-4	14	Candesartan	Candesartan	9	✓
	Olmesartan	2	25.6	Olmesartan	Olmesartan	10–15	✓
	Valsartan	2-4	23	Valsartan	Valsartan	9	✓
Uricostatic	Allopurinol	1.5	67	Oxypurinol	(a) Allopurinol (b) Oxypurinol	18–43	✓
Antihyperglycemic	Metformin	2.5	30-60	Excreted unchanged	Metformin	5	✓
	Sitagliptin	1–4	87	Sitagliptin	Sitagliptin	12	✓
Antiepileptic	Gabapentin	2–3	60	No metabolization	Gabapentin	5–7	✓
Antidepressant	Citalopram	3	80	Desmethylcitalopram	(a) Citalopram-*N*-oxide (b) Citalopram propionate (c) Desmethylcitalopram (d) Escitalopram	35	✓ Combined with escitalopram
	Escitalopram	3–4	80	Desmethylcitalopram	(a) Citalopram-*N*-oxide (b) Citalopram propionate (c) Desmethylcitalopram (d) Escitalopram	35	✓ Combined with citalopram
	Fluoxetine	6–8	72	Norfluoxetine	Fluoxetine	96–144	✓
	Venlafaxine	2–3	40–45	O-desmethylvenlafaxine	(a) Venlafaxine (b) *O*-desmethylvenlafaxine	5–11	✓
Antipsychotic	Quetiapine	1	100	N-desalkylquetiapine	(a) Quetiapine (b) *N*-desalkylquetiapine	7–12	✓

whether their pharmacokinetic properties were suitable for validating self-reported drug use (Table 1), and we excluded enalapril after this assessment. The panel measures enalapril, rather than the active metabolite enalaprilat. Since enalapril is a rapidly transformed prodrug (plasma elimination $t_{½}$ < 1 hour), we expected many false positives. Based on our pharmacokinetic assessment, we also combined citalopram and escitalopram. Escitalopram is the therapeutically active *S*-enantiomer of the racemic mixture citalopram, therefore differentiation between the intake of both drugs via the metabolomic panel is not possible. Ultimately, we analysed 12 individual drugs and the combination of citalopram and escitalopram (Table 1).

2.6 | Study population

This cross-sectional study was conducted on the first 5000 participants of the Rhineland Study who completed their study visits between March 2016 and February 2020. Analyses were based on 4500 participants for whom we measured drug metabolites in plasma. Of these, we had to exclude 114 participants because (i) metabolomics data were incomplete (n = 29), (ii) self-reported medication data were incomplete (n = 71) and (iii) use of the respective drug was reported as needed (n = 14). The analytical sample comprised 4386 individuals (women, n = 2459; men, n = 1927; <65 years, n = 3216; ≥65 years, n = 1170).

2.7 | Statistical analysis

Descriptive statistics are presented for women and men. Groups were compared using chi-square tests (categorical variables) and ANOVA tests (continuous variables). We assessed concordance between self-reported medication intake and drug metabolites measured in plasma. Participants' self-report data were classified as (i) *true positive* (metabolite detected and self-reported use), (ii) *true negative* (metabolite not detected and no self-reported use), (iii) *false positive* (metabolite not detected but self-reported use) or (iv) *false negative* (metabolite detected but no self-reported use). Concordances between self-reported drug use and measured metabolites were calculated using Cohen's kappa (κ). The kappa statistic measures the extent of inter-rater agreement. Kappa values were classified as (i) ≤0.4 *poor to fair agreement*, (ii) 0.41–0.6 *moderate agreement*, (iii) 0.61–0.8 *substantial agreement* and (iv) >0.8 *almost perfect agreement*.[33] All κ values are presented with 95% confidence intervals (CI). Additional analyses were performed stratified for sex and age (<65 *vs* ≥65 years), when the number of users in each subgroup was ≥15. All statistical analyses were performed using R version 3.6.1.

2.8 | Nomenclature of targets and ligands

Key protein targets and ligands in this article are hyperlinked to corresponding entries in http://www.guidetopharmacology.org, the common portal for data from the IUPHAR/BPS Guide to PHARMACOLOGY, and are permanently archived in the Concise Guide to PHARMACOLOGY 2019/20.[34]

3 | RESULTS

3.1 | General characteristics of the study population

The characteristics of the analytical sample of 4386 individuals are presented in Table 2. Among included participants, the mean age was 55.0 years (standard deviation [SD] 14.0 years, range 30–95) and 26.7% (n = 1170) of the individuals were ≥65 years. Included participants were younger (P < .001) than those excluded (mean age 59.9 years, SD 13.7 years, range 30–87). The ratio of women to men was higher (P < .001) in included participants (56.1% women, n = 2459) compared to those excluded (45.6% women, n = 52). On average, participants used 2.1 drugs (SD 2.6 drugs) and individuals aged ≥65 years (3.8 drugs, SD 3.3) used significantly (P < .001) more drugs than those aged <65 years (1.7 drugs, SD 2.1). Elevan of the 13 drugs analysed are among the top 40 most frequently, regularly used drugs in the Rhineland Study (Table A3).

3.2 | Accuracy of self-reported medication use

Of the 13 drugs analysed, 11 showed an almost perfect agreement (κ > 0.8) between self-reported medication use and metabolites measured in plasma (Table 3). For sitagliptin, the agreement was only slightly lower, with κ = 0.80 (46 users, 95% CI 0.71–0.90). For hydrochlorothiazide, the concordance was only 0.61 (355 users, 95% CI 0.56–0.66), as we did not detect metabolites in 187 self-reported users.

Concordance rates for antihypertensive drugs other than hydrochlorothiazide were comparable to concordance rates for the included antidepressants and quetiapine (Table 3). The concordance rates of antihyperglycemic medications tended to be lower (κ < 0.90) than those of antihypertensives, antidepressants and quetiapine, but still showed almost perfect/substantial concordance.

We performed sex-stratified analyses for drugs from each included drug class, except for gabapentin (Figure 1). For five of eight of these drugs, concordance was higher for women than men. There was no pattern with respect to drug classes. A statistically significant difference between men and women was observed only for allopurinol, with higher concordance in men (women, 15 users, κ = 0.68, 95% CI,0.52–0.83; men, 77 users, κ = 0.89, 95% CI 0.84–0.95). Detailed results on concordance rates stratified by sex can be found in Table A4.

We performed age-stratified analyses for drugs from each included drug class, except for gabapentin (Figure 2). Concordance was higher for seven of nine drugs in individuals ≥65 years compared to individuals aged <65 years. However, in all cases the 95% CIs of

TABLE 2 Descriptive characteristics of the population: total sample and stratified for sex

	Total sample	Missing (%)	Women	Men	*P* value
Number of participants, n (%)	4386		2459 (56.1)	1927 (43,9)	<.001
Age in years, M (SD)	55.0 (14.0)	0.0	54.8 (13.7)	55.2 (14.4)	.315
Age groups		0.0			.069
< 65 years	3216 (73.3)		1830 (74.4)	1386 (71.9)	
≥ 65 years	1170 (26.7)		629 (25.6)	541 (28.1)	
Education, n (%)		0.8			<.001
Low	81 (1.9)		63 (2.6)	18 (0.9)	
Middle	1950 (44.8)		1214 (49.8)	736 (38.5)	
High	2319 (53.3)		1159 (47.6)	1160 (60.6)	
Smoking, n (%)		4.0			<.001
Never	1992 (47.3)		1186 (50.3)	806 (43.5)	
Former	1701 (40.4)		901 (38.2)	800 (43.2)	
Current	519 (12.3)		273 (11.6)	246 (13.3)	
BMI (kg/m^2), M (SD)	25.9 (4.5)	0.5	25.4 (4.8)	26.5 (4.0)	<.001
Diabetes, n (%)	223 (5.1)	0.8	88 (3.6)	135 (7.0)	<.001
Hypertension, n (%)	1625 (37.6)	1.4	808 (33.4)	817 (42.8)	<.001
Hypercholesterolemia, n (%)	1693 (38.9)	0.8	944 (38.7)	749 (39.2)	.779
Polypharmacy, n (%)					<.001
All	582 (13.3)	0.0	340 (13.8)	242 (12.6)	
< 65 years	231 (7.2)		142 (7.8)	89 (6.4)	
≥ 65 years	351 (30.0)		198 (31.5)	153 (28.3)	
Average number of prescribed drugs, M (SD)		0.0			<.001
All	2.1 (2.6)		2.3 (2.6)	1.8 (2.7)	
< 65 years	1.4 (2.0)		1.7 (2.1)	1.1 (1.8)	
≥ 65 years	3.8 (3.3)		3.9 (3.1)	3.8 (3.5)	

Note: Education based on International Standard Classification of Education (low, lower secondary education or below; middle, upper secondary education to undergraduate university level; high, postgraduate university study). Diabetes based on current antidiabetic drug use/glycated haemoglobin (Hba1c) (no diabetes <6.5%, diabetes ≥6.5%) measured in fasting morning blood.
Hypertension based on measured mean systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg, and/or antihypertensive drug use (irrespective of blood pressure levels). Hypercholesterolemia based on self-report. Polypharmacy defined as regular use of ≥5 prescribed drugs and supplements.
Abbreviations: BMI, body mass index; M, mean; n, number; SD, standard deviation.

the κ values overlapped, indicating no significant age-related differences. Detailed results on concordance rates stratified by age can be found in Table A5.

4 | DISCUSSION

We validated self-reported medication intake with drug metabolites measured in plasma with untargeted mass spectrometry-based metabolomics techniques in a general population (≥30 years). We included 13 drugs commonly used to treat hypertension, diabetes, depression, gout and epilepsy. We found almost perfect agreement rates (κ > 0.8) for 11 out of the 13 drugs analysed. Although psychoactive drugs are considered to be prone to stigmatization bias, we found an almost perfect agreement for all included antidepressants and quetiapine with concordances of κ > 0.9. Allopurinol was the only drug for which we observed significant sex differences, with higher concordance in men compared to women. There were no significant differences in concordance rates between younger and older age groups. The high concordance between self-reported medication use and measured drug metabolites suggests that interview-based self-reported medication data is a reliable and accurate method to assess regular drug intake, regardless of drug class, age and sex.

The diuretic hydrochlorothiazide showed the lowest concordance ($\kappa = 0.61$) between self-reported drug intake and measured metabolites. This was primarily due to a high number of false positives (187 of 355 regular self-reported users), meaning individuals reported taking hydrochlorothiazide but no metabolites were detected. Frequent urination is often a consequence of taking diuretics. Therefore, a possible explanation may be that participants did not take their diuretic when going out for the day to participate in our study. We do

TABLE 3 Concordance between self-reported drug use and measured metabolites with kappa (κ) values (95% CI)

Drug class	Drug	Regular users	TP	FP	TN	FN	Cohen's kappa (κ) (95% CI)
Antihypertensive	Hydrochlorothiazide	355	168	187	4023	8	0.61 (0.56–0.66)
	Metoprolol	101	92	9	4265	20	**0.86 (0.81–0.91)**
	Candesartan	279	269	10	4103	4	**0.97 (0.96–0.99)**
	Olmesartan	38	35	3	4348	0	**0.96 (0.91–1.00)**
	Valsartan	98	94	4	4284	4	**0.96 (0.93–0.99)**
Uricostatic	Allopurinol	92	85	7	4271	23	**0.85 (0.79–0.90)**
Antihyperglycemic	Metformin	109	89	20	4274	3	**0.88 (0.84–0.93)**
	Sitagliptin	46	31	15	4340	0	0.80 (0.71–0.90)
Antiepileptic	Gabapentin	19	14	5	4366	1	**0.82 (0.68–0.96)**
Antidepressant	(Es-) citalopram	90	86	4	4283	13	**0.91 (0.86–0.95)**
	Fluoxetine	19	19	0	4363	4	**0.90 (0.81–1.00)**
	Venlafaxine	49	47	2	4334	3	**0.95 (0.90–0.99)**
Antipsychotic	Quetiapine	16	16	0	4368	2	**0.94 (0.86–1.00)**

Note: Kappa values in bold show an almost perfect agreement (κ > 0.8).
Abbreviations: CI, confidence interval; FN, false negative (metabolite detected but no self-reported use); FP, false positive (metabolite not detected but self-reported use); TN, true negative (metabolite not detected and no self-reported use); TP, true positive (metabolite detected and self-reported use).

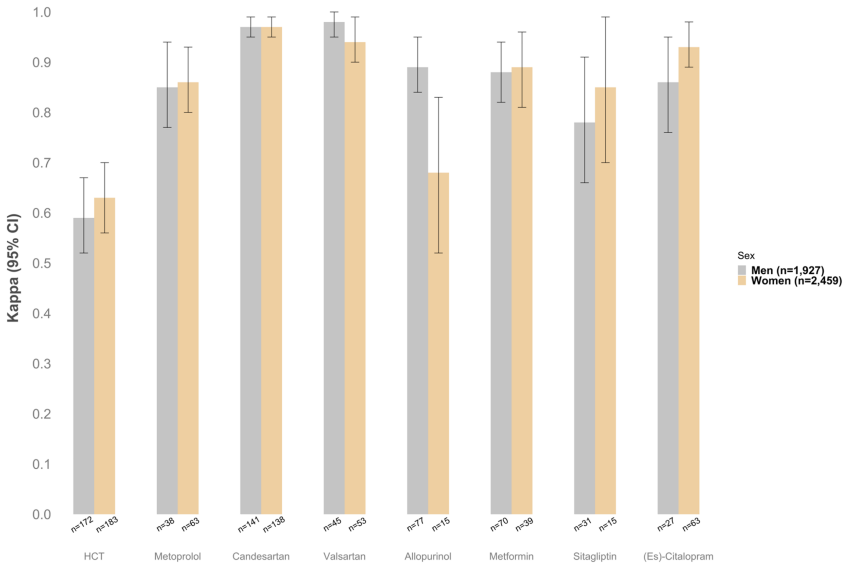


FIGURE 1 κ values with 95% CI of self-reported drug use *vs* drug metabolites measured in plasma, stratified for sex. n, number of drug users

not anticipate that hydrochlorothiazide degradation played a role in our samples because we detected hydrochlorothiazide in the longest-stored (>21 months) samples from low-dose users.

The concordance rate for sitagliptin was $\kappa = 0.80$. This borderline perfect agreement (κ > 0.8) was due to 15 false positives who reported medication intake but the drug metabolites were not detected. Because blood samples were collected before breakfast, it is likely that the last sitagliptin intake of most participants was more than 24 hours ago. Considering the elimination $t_{½}$ is 12 hours, this could explain why sitagliptin was not detected in some individuals.

Differences in concordances between older and younger age groups were small and nonsignificant, but generally higher in older age groups (≥65 years) compared to younger individuals (<65 years). This is consistent with the findings of Sediq et al[12] and contradicts the

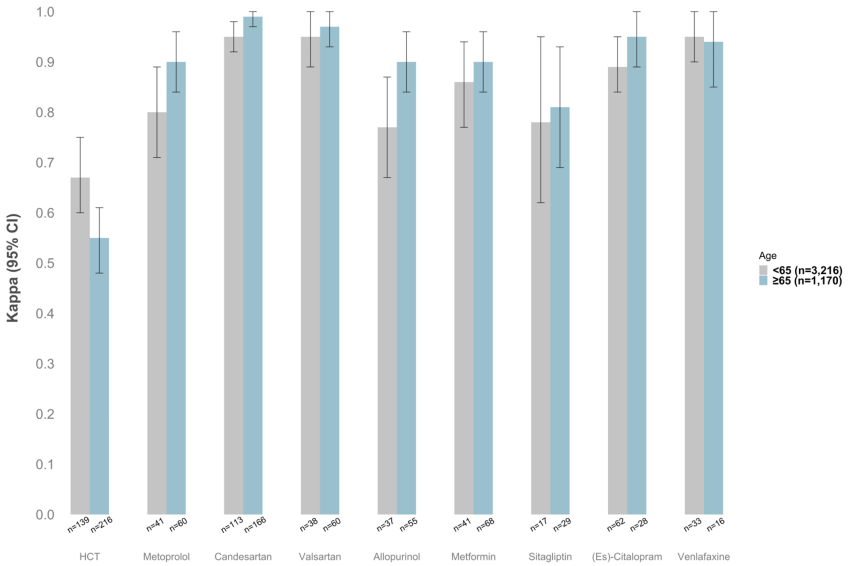


FIGURE 2 κ values with 95% CI of self-reported drug use *vs* drug metabolites measured in plasma, stratified for age. n, number of drug users

common assumption that the ability to recall medication use decreases with age as a result of cognitive decline and higher and more complex medication use.[35,36] Consistent with previous studies, we observed no effect of sex on concordance rates; we found only a significantly higher concordance in men compared to women for allopurinol.[11,12,16] However, there was a large difference in the prevalence of allopurinol use, with 77 users in men versus 15 in women.

Concerns regarding the accuracy of self-reported medication use mostly relate to the reliability of reporting certain drug classes vulnerable to stigmatization bias, such as antidepressants. While most studies validating self-reported antidepressant use with EMRs found concordance below 0.63,[10–12] some also reported good concordance.[13] In our study, we were able to analyse the validity of four antidepressants (citalopram/escitalopram, fluoxetine, venlafaxine) and one antipsychotic drug (quetiapine). For all of these drugs, we found concordance rates of ≥0.9, which may indicate that (i) the self-reported data collected in our study are of better quality than self-reported data from other studies and/or (ii) self-reported medication use better reflects actual intake than EMRs. It has been reported that more than 25% of individuals who are prescribed antidepressants for the first time decline treatment; they either do not start treatment or do not persist with antidepressant use for more than 2 weeks, which may result in inaccurate EMRs.[37]

We also found generally similar or even higher concordance rates compared with previous studies comparing self-reported medication use with EMRs.[10–12,16,19] As mentioned before, this may suggest that lower concordances are not due to inaccuracies in self-reported medication use but to inaccuracies in EMRs. Studies comparing multiple drug metabolites measured in blood using LC–MS/MS techniques with EMRs reported high rates of discrepancies between the two.[21,22] Even antihypertensive drugs, which are not prone to misuse, were not prescribed in 14–26% of cases where they were detected, suggesting that EMRs can be incomplete.[22] This can also be deduced from findings of a Dutch study comparing self-reported medication use with prescription records that unexpectedly found overreporting, ie, the drug was self-reported but not listed in prescription records.[12]

The strength of our study is that we validated self-reported medication use with metabolites measured in plasma, which is more reflective of actual drug use than EMRs. Furthermore, our data on self-reported medication use were assessed in a manner comparable to the brown bag method. This method is designed to provide a more complete overall picture of an individual's current medication profile compared to pharmacy records.[38,39] However, it is unclear whether self-reported data based on self-administered questionnaires, which are less time-consuming, are of the same quality. Another strength is that we were able to assess concordances for drugs prone to stigmatization bias, such as antidepressants, as well as for drugs not susceptible to stigma, such as antihypertensives. Although we were able to assess concordance rates for drug classes at high risk for stigmatization bias, we were unable to assess medication classes often dispensed in hospitals. For drugs dispensed in hospitals, EMRs might not be a good source for assessing exposure, and the quality of self-reported intake of those drugs is unclear. A limitation of our study is that we could not select the drugs according to their frequency of use. Nevertheless, 11 of the 13 drugs analysed were among the top 40 most frequently, regularly used drugs in the Rhineland Study.

Another limitation of our study is that the accuracy of self-reported medication use may be culture-dependent, particularly for drugs vulnerable to stigma. In addition, our study population could represent a "healthier" population, as health status could influence drug metabolism, which could limit the generalizability of the results. Although frail older adults are less likely to participate in our study, our population does not appear to be healthier than the general German population, as, for example, the prevalence of hypertension and polypharmacy is comparable.[40,41] Furthermore, we used an untargeted metabolomics panel, which cannot detect all marketed drugs. Nevertheless, we were able to analyse concordances of a variety of commonly used drugs from different drug classes. Although an untargeted metabolomic approach does not allow quantification but rather focuses on simultaneous detection of different metabolites, we do not consider the use of an untargeted approach to be critical, as we only assessed drug use (metabolite detected) and nonuse (metabolite not detected), and were not interested in actual concentrations. Although, to our knowledge, this is the largest study on concordance rates of self-reported medication use and measured metabolites in plasma, for some drugs the number of individuals using the respective drug was still small, resulting in wide confidence intervals. A small number of users may be the reason why we found no indications of age- or sex-related effects on concordances. Another limitation could be that we did not adapt the blood sampling and storage to the individual characteristics of the specific drug metabolite. To ensure that included drug metabolites were suitable for detection, we checked the pharmacokinetic properties of the drugs a priori. Finally, variability in longevity and detection of drug metabolites may depend on age, health status and genetic, metabolic and microbiome variations.[42–45]

Our study shows that in our general population self-reported medication intake is a reliable and accurate method to assess medication intake, even for drugs considered to be prone to stigmatization.

ACKNOWLEDGEMENTS

This work was supported by a grant provided by the Federal Institute for Drugs and Medical Devices (BfArM), Bonn, Germany (V-15646/68502/2014–2016 and V-17746/68502/2018–2020). The funding sources have not been involved in study design; in the collection, analysis and interpretation of the data, in the writing or in the decision to submit the article for publication.

COMPETING INTERESTS

The authors report no competing interests in this work.

CONTRIBUTORS

N.A. conceptualization, methodology, formal analysis, writing- original draft, visualization. J.C.S. conceptualization, methodology, writing – review and editing, supervision. M.M.B.B. conceptualization, methodology, resources, data curation, writing – review and editing, supervision, funding acquisition. F.M.deV. conceptualization, methodology, data curation, writing – review and editing, visualization, supervision, project administration.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Nersi Alaeddin https://orcid.org/0000-0003-1960-9812
Julia C. Stingl https://orcid.org/0000-0002-1566-8156
Monique M. B. Breteler https://orcid.org/0000-0002-0626-9305
Folgerdiena M. de Vries https://orcid.org/0000-0002-8621-1532

REFERENCES

1. Reijneveld SA. The cross-cultural validity of self-reported use of health care: A comparison of survey and registration data. *J Clin Epidemiol*. 2000;53(3):267-272. doi:10.1016/S0895-4356(99)00138-9
2. Peersman W, Pasteels I, Cambier D, De Maeseneer J, Willems S. Validity of self-reported utilization of physician services: a population study. *Eur J Public Health*. 2014;24(1):91-97. doi:10.1093/eurpub/ckt079
3. Skurtveit S, Selmer R, Tverdal A, Furu K. The validity of self-reported prescription medication use among adolescents varied by therapeutic class. *J Clin Epidemiol*. 2008;61(7):714-717. doi:10.1016/j.jclinepi.2007.11.013
4. Klungel OH, De Boer A, Paes AHP, Herings RMC, Seidell JC, Bakker A. Influence of question structure on the recall of self-reported drug use. *J Clin Epidemiol*. 2000;53(3):273-277. doi:10.1016/S0895-4356(99)00167-5
5. Cotterchio M, Kreiger N, Darlington G, Steingart A. Comparison of self-reported and physician-reported antidepressant medication use. *Ann Epidemiol*. 1999;9(5):283-289. doi:10.1016/S1047-2797(98)00072-6
6. Knudsen P, Hansen EH, Traulsen JM. Perceptions of young women using SSRI antidepressants — a reclassification of stigma. *Int J Pharm Pract*. 2011;10(4):243-252. doi:10.1211/096176702776868415
7. Goldberg EM, Marks SJ, Merchant RC, Nagy JL, Aquilante JA, Beaudoin FL. How accurately do older adult emergency department patients recall their medications? *Acad Emerg Med*. 2021;28(2):248-252. doi:10.1111/acem.14032
8. Jones GT, Tabassum V, Zarow GJ, Ala TA. The inability of older adults to recall their drugs and medical conditions. *Drugs Aging*. 2015;32(4):329-336. doi:10.1007/s40266-015-0255-z
9. West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: Self-report compared with database information. *Am J Epidemiol*. 1995;142(10):1103-1112. doi:10.1093/oxfordjournals.aje.a117563
10. Nielsen MW, Søndergaard B, Kjøller M, Hansen EH. Agreement between self-reported data on medicine use and prescription records vary according to method of analysis and therapeutic group. *J Clin Epidemiol*. 2008;61(9):919-924. doi:10.1016/j.jclinepi.2007.10.021
11. Richardson K, Kenny RA, Peklar J, Bennett K. Agreement between patient interview data on prescription medication use and pharmacy records in those aged older than 50 years varied by therapeutic group and reporting of indicated health conditions. *J Clin Epidemiol*. 2013;66(11):1308-1316. doi:10.1016/j.jclinepi.2013.02.016
12. Sediq R, van der Schans J, Dotinga A, et al. Concordance assessment of self-reported medication use in the netherlands three-generation lifelines cohort study with the pharmacy database IADB.nl: The pharmlines initiative. *Clin Epidemiol*. 2018;10:981-989. doi:10.2147/CLEP.S163037.
13. Hafferty JD, Campbell AI, Navrady LB, et al. Self-reported medication use validated through record linkage to national

prescribing data. *J Clin Epidemiol*. 2018;94:132-142. doi:10.1016/j.jclinepi.2017.10.013
14. Klungel OH, De Boer A, Paes AHP, Herings RMC, Seidell JC, Bakker A. Agreement between self-reported antihypertensive drug use and pharmacy records in a population-based study in The Netherlands. *Pharm World Sci*. 1999;21:217-220. doi:10.1023/A:1008741321384
15. Haukka J, Suvisaari J, Tuulio-Henriksson A, Lönnqvist J. High concordance between self-reported medication and official prescription database information. *Eur J Clin Pharmacol*. 2007;63(11):1069-1074. doi:10.1007/s00228-007-0349-6
16. Sjahid SI, Van Der Linden PD, Stricker BHC. Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: The Rotterdam elderly study. *Br J Clin Pharmacol*. 1998;45(6):591-595. doi:10.1046/j.1365-2125.1998.00716.x
17. Wilke T, Müller S, Morisky DE. Toward identifying the causes and combinations of causes increasing the risks of nonadherence to medical regimens: Combined results of two German self-report surveys. *Value Health*. 2011;14(8):1092-1100. doi:10.1016/j.jval.2011.06.002
18. Noize P, Bazin F, Dufouil C, et al. Comparison of health insurance claims and patient interviews in assessing drug use: Data from the Three-City (3C) Study. *Pharmacoepidemiol Drug Saf*. 2009;18(4):310-319. doi:10.1002/pds.1717
19. Kwon A, Bungay KM, Pei Y, et al. Antidepressant use: Concordance between self-report and claims records. *Med Care*. 2003;41(3):368-374. doi:10.1097/00005650-200303000-00005
20. Løkkegaard EL, Johnsen SP, Heitmann BL, et al. The validity of self-reported use of hormone replacement therapy among Danish nurses. *Acta Obstet Gynecol Scand*. 2004;83(5):476-481. doi:10.1111/j.0001-6349.2004.00376.x
21. Ryan TP, Morrison RD, Sutherland JJ, et al. Medication adherence, medical record accuracy, and medication exposure in real-world patients using comprehensive medication monitoring. *PLoS ONE*. 2017;12(9):1-19. doi:10.1371/journal.pone.0185471
22. Sutherland JJ, Morrison RD, McNaughton CD, et al. Assessment of patient medication adherence, medical record accuracy, and medication blood concentrations for prescription and over-the-counter medications. *JAMA Netw Open*. 2018;1(7):e184196. doi:10.1001/jamanetworkopen.2018.4196
23. Dennis K, Carter B, Gapstur S, Stevens V. Metabolomics approach for validation of self-reported ibuprofen and acetaminophen use. *Metabolites*. 2018;8(4):55. doi:10.3390/metabo8040055
24. Vries FM, Stingl JC, Breteler MMB. Polypharmacy, potentially inappropriate medication and pharmacogenomics drug exposure in the Rhineland Study. *Br J Clin Pharmacol*. 2021;87(7):2732-2756. doi:10.1111/bcp.14671
25. Mühlberger N, Behrend C, Stark RHR. Database-supported indentification and entry of drug data in health studies – experience with the IDOM software. *Inform Biometrie Und Epidemiol Medizin Und Biol*. 2003;34:601-611.
26. Wissenschaftliches Institut der AOK. n.d. www.wido.de (accessed January 17, 2021).
27. ATC/DDD index. n.d. https://www.whocc.no/atc_ddd_index/ (accessed January 17, 2021).
28. Evans AM, DeHaven CD, Barrett T, Mitchell M, Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal Chem*. 2009;81(16):6656-6667. doi:10.1021/ac901536h
29. Suhre K, Shin SY, Petersen AK. Human metabolic individuality in biomedical and pharmaceutical research. *Nature*. 2011;477:54-60. doi:10.1038/nature10354
30. Shin S, Fauman EB, Petersen A, et al. Atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014;46(6):543-550. doi:10.1038/ng.2982.An
31. Dehaven CD, Evans AM, Dai H, Lawton KA. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J Chem*. 2010;2(1):1-12. doi:10.1186/1758-2946-2-9
32. Ford L, Kennedy AD, Goodman KD, et al. Precision of a clinical metabolomics profiling platform for use in the identification of inborn errors of metabolism. *J Appl Lab Med*. 2020;5(2):342-356. doi:10.1093/jalm/jfz026
33. Viera AJ, Garrett JM. Understanding interobserver agreement: The kappa statistic. *Fam Med*. 2005;37(5):360-363.
34. Alexander SPH, Kelly E, Mathie A, et al. The Concise Guide to Pharmacology 2019/20: Transporters. *Br J Pharmacol*. 2019;176(S1):S397-S493. doi:10.1111/bph.14753
35. Kantor ED, Rehm CD, Haas JS, Chan AT, Giovannucci EL. Trends in prescription drug use among adults in the United States from 1999-2012. *J Am Med Assoc*. 2015;314(17):1818-1831. doi:10.1001/jama.2015.13766
36. Welker KL, Mycyk MB. Pharmacology in the Geriatric Patient. *Emerg Med Clin North Am*. 2016;34(3):469-481. doi:10.1016/j.emc.2016.04.003
37. van Geffen EC, Gardarsdottir H, van Hulten R, van Dijk L, Egberts AC, Heerdink ER. Initiation of antidepressant therapy: do patients follow the GP's prescription? *Br J Gen Pract*. 2009;59(559):81-87. doi:10.3399/bjgp09X395067
38. Nathan A, Goodyer L, Lovejoy A, Rashid A. "Brown bag" medication reviews as a means of optimizing patients' use of medication and of identifying potential clinical problems. *Fam Pract*. 1999;16(3):278-282. doi:10.1093/fampra/16.3.278
39. Caskie GIL, Willis S, Warner Schaie K, Zanjani F. Congruence of medication information from a brown bag data collection and pharmacy records: Findings from the Seattle Longitudinal Study. *Exp Aging Res*. 2006;32(1):79-103. doi:10.1080/03610730500326341
40. Holstiege J, Akmatov MK, Steffen A & Bätzing J Diagnosis prevalence of hypertension in Germany – recent findings from nation-wide ambulatory health care 2020. 10.20364/VA-20.01
41. Moßhammer D, Haumann H, Mörike K, Joos S. Polypharmacy – An upward trend with unpredictable effects. *Dtsch Arztebl Int*. 2016;113:627-633. doi:10.3238/arztebl.2016.0627
42. Javdan B, Lopez JG, Chankhamjon P, et al. Personalized mapping of drug metabolism by the human gut microbiome. *Cell*. 2020;181(7):1661-1679.e22. doi:10.1016/j.cell.2020.05.001
43. Mangoni AA, Jackson SHD. Age-related changes in pharmacokinetics and pharmacodynamics: Basic principles and practical applications. *Br J Clin Pharmacol*. 2003;57(1):6-14. doi:10.1046/j.1365-2125.2003.02007.x
44. Klaassen CD, Cui JY. Review: Mechanisms of how the intestinal microbiota alters the effects of drugs and bile acids. *Drug Metab Dispos*. 2015;43(10):1505-1521. doi:10.1124/dmd.115.065698
45. Pinto N, Dolan ME. Clinically relevant genetic variations in drug metabolizing enzymes. *Curr Drug Metab*. 2011;12(5):487-497. doi:10.2174/138920011795495321

How to cite this article: Alaeddin N, Stingl JC, Breteler MMB, de Vries FM. Validation of self-reported medication use applying untargeted mass spectrometry-based metabolomics techniques in the Rhineland study. *Br J Clin Pharmacol*. 2022;88(5):2380-2395. doi:10.1111/bcp.15175

APPENDIX A

TABLE A1 Drugs and corresponding metabolites in the metabolomics panel per drug class

Drug class	Drug	Measured metabolite(s)
Analgesic/anesthetic	Acetaminophen	(a) 2-acetamidophenol sulphate (b) 2-hydroxyacetaminophen sulphate (c) 2-methoxyacetaminophen glucuronide (d) 2-methoxyacetaminophen sulphate (e) 3-(cystein-*S*-yl) acetaminophen (f) 3-(methylthio) acetaminophen sulphate (g) 3-(*N*-acetyl-L-cystein-*S*-yl) acetaminophen (h) 4-acetamidophenol (i) 4-acetamidophenylglucuronide (j) 4-acetaminophen sulphate
	Celecoxib	Celecoxib
	Ibuprofen	(a) Ibuprofen (b) 2-hydroxyibuprofen (c) Carboxyibuprofen (d) Carboxyibuprofen glucuronide (e) Ibuprofen acyl glucuronide
	Naproxen	(a) Naproxen (b) Desmethylnaproxen (c) Desmethylnaproxen sulfate
	Tramadol	(a) Tramadol (b) *N*-desmethyl tramadol (c) *O*-desmethyltramadol (d) *O*-desmethyltramadol glucuronide
Antibiotic	Doxycycline	Doxycycline
	Fluconazole	Fluconazole
	Ofloxacin	Ofloxacin
Anti-inflammatory	*N*-acetyl sulfapyridine	*N*-acetyl sulfapyridine
Antimalarial	Quinine	Quinine
Immunosuppressant	Mycophenolic acid	(a) Mycophenolic acid (b) Mycophenolic acid glucuronide
Antihypertensive	Enalapril	Enalapril
	Hydrochlorothiazide	Hydrochlorothiazide
	Metoprolol	(a) Metoprolol (b) Metoprolol acid
	Candesartan	Candesartan
	Olmesartan	Olmesartan
	Valsartan	Valsartan
Antiarrhythmic	Verapamil	Verapamil
Ulcer therapeutic	Omeprazole	Omeprazole
	Pantoprazole	Pantoprazole
	Ranitidine	(a) Ranitidine (b) Ranitidine *N*-oxide
Uricostatic	Allopurinol	(a) Allopurinol (b) Oxypurinol
Antihyperglycemic	Metformin	Metformin
	Sitagliptin	Sitagliptin
Antiepileptic	Carbamazepine	(a) Carbamazepine (b) Carbamazepine 10,11-epoxide (c) Carbamazepine glucuronide

TABLE A1 (Continued)

Drug class	Drug	Measured metabolite(s)
	Gabapentin	Gabapentin
	Lamotrigine	Lamotrigine
	Levetiracetam	Levetiracetam
	Pregabalin	Pregabalin
	Topiramate	Topiramate
	Valproate	(a) Valproate (b) 2-propyl-4-pentenoate (4-ene-valproate) (c) 3-hydroxyvalproate
Antidepressant	Citalopram	(a) Citalopram-*N*-oxide (b) Citalopram propionate (c) Desmethylcitalopram (d) Escitalopram
	Escitalopram	(a) Citalopram-*N*-oxide (b) Citalopram propionate (c) Desmethylcitalopram (d) Escitalopram
	Duloxetine	4-hydroxy duloxetine glucuronide
	Fluoxetine	Fluoxetine
	Venlafaxine	(a) Venlafaxine (b) *O*-desmethylvenlafaxine
Antipsychotic	Quetiapine	(a) Quetiapine (b) *N*-desalkylquetiapine
Analgesic/antiemetic	Tetrahydrocannabinol	(a) Tetrahydrocannabinol carboxylic acid (b) Tetrahydrocannabinol carboxylic acid glucuronide
Antihistamine	Cetirizine	Cetirizine
	Diphenhydramine	Diphenhydramine
	Fexofenadine	Fexofenadine
Antitussive	Dextromethorphan	Dextromethorphan
Topic agent	Hydroquinone sulphate	Hydroquinone sulphate
	Salicylate	Salicylate

TABLE A2 Retention index and mass-to-charge ratio (*m*/*z*) for analysed drugs

Drug	Measured metabolite(s)	Retention index	Mass-to-charge ratio (*m*/*z*)
Enalapril	Enalapril	900	377.2071
Hydrochlorothiazide	Hydrochlorothiazide	2046	295.9572
Metoprolol	(a) Metoprolol	(a) 760	(a) 268.1907
	(b) Metoprolol acid	(b) 2950	(b) 268.1543
Candesartan	Candesartan	3436	439.1524
Olmesartan	Olmesartan	3747	445.1994
Valsartan	Valsartan	4228	434.2198
Allopurinol	(a) Allopurinol	(a) 1530	(a) 135.0312
	(b) Oxypurinol	(b) 1656	(b) 151.0262
Metformin	Metformin	2817	130.1087
Sitagliptin	Sitagliptin	4485	468.1112
Gabapentin	Gabapentin	2407.4	170.1187
Citalopram	(a) Citalopram-*N*-oxide	(a) 905	(a) 341.166
	(b) Citalopram propionate	(b) 4400	(b) 310.0885
	(c) Desmethylcitalopram	(c) 880	(c) 311.1554
	(d) Escitalopram	(d) 908	(d) 325.1711
Escitalopram	(a) Citalopram-*N*-oxide	(a) 905	(a) 341.166
	(b) Citalopram propionate	(b) 4400	(b) 310.0885
	(c) Desmethylcitalopram	(c) 880	(c) 311.1554
	(d) Escitalopram	(d) 908	(d) 325.1711
Fluoxetine	Fluoxetine	1025	310.1413
Venlafaxine	(a) Venlafaxine	(a) 848	(a) 278.2115
	(b) *O*-desmethylvenlafaxine	(b) 660	(b) 264.1958
Quetiapine	(a) Quetiapine	(a) 874	(a) 384.174
	(b) *N*-desalkylquetiapine	(b) 874	(b) 296.1958

TABLE A3 Top 50 drugs used regularly in the Rhineland Study

Rank	Drug	Rank	Drug
1	Levothyroxine	26	Losartan
2	Acetylsalicylic acid	27	**Venlafaxine**
3	**Hydrochlorothiazide**	28	**Citalopram**
4	Bisoprolol	29	Ezetimibe
5	**Candesartan**	30	**Sitagliptin**
6	Ramipril	31	**Enalapril**
7	Atorvastatin	32	Rivaroxaban
8	Amlodipine	33	Clopidogrel
9	Pantoprazole	34	Desogestrel
10	Simvastatin	35	Prednisolone
11	Estradiol	36	Omeprazole
12	Ethinylestradiol	37	**Olmesartan**
13	**Metformin**	38	**Escitalopram**
14	**Metoprolol**	39	Phenprocoumon
15	**Valsartan**	40	Amitriptyline
16	Formoterol	41	Levonorgestrel
17	Estriol	42	Telmisartan
18	**Allopurinol**	43	Nebivolol
19	Tamsulosin	44	Mirtazapine
20	Budesonide	45	Irbesartan
21	Progesterone	46	Lisinopril
22	Latanoprost	47	Lercanidipine
23	Beclomethasone	48	Fluticasone
24	Torasemide	49	Ibuprofen
25	Dienogest	50	Levodopa

Note. Analysed drugs in bold

TABLE A4 Concordance between self-reported drug use and measured metabolites with kappa (κ) values (95% CI) stratified for sex

		Women (n = 2459)						Men (n = 1927)					
Drug class	**Drug**	**Regular users**	**TP**	**FP**	**TN**	**FN**	**Cohen's kappa (κ) (95% CI)**	**Regular users**	**TP**	**FP**	**TN**	**FN**	**Cohen's kappa (κ) (95% CI)**
Antihypertensive	Hydrochlorothiazide	183	90	93	2271	5	0.63 (0.56–0.7)	172	78	94	1752	3	0.59 (0.52–0.67)
	Metoprolol	63	56	7	2386	10	**0.86 (0.80–0.93)**	38	36	2	1879	10	**0.85 (0.77–0.94)**
	Candesartan	138	133	5	2319	2	**0.97 (0.95–0.99)**	141	136	5	1784	2	**0.97 (0.95–0.99)**
	Olmesartan	13	11	2	2446	0	**0.92 (0.8–1.00)**	25	24	1	1902	0	**0.98 (0.94–1.00)**
	Valsartan	53	50	3	2403	3	**0.94 (0.90–0.99)**	45	44	1	1881	1	**0.98 (0.95–1.00)**
Uricostatic	Allopurinol	15	15	0	2430	14	0.68 (0.52–0.83)	77	70	7	1841	9	**0.89 (0.84–0.95)**
Antihyperglycemic	Metformin	39	32	7	2419	1	**0.89 (0.81–0.96)**	70	57	13	1855	2	**0.88 (0.82–0.94)**
	Sitagliptin	15	11	4	2444	0	**0.85 (0.70–0.99)**	31	20	11	1896	0	0.78 (0.66–0.91)
Antiepileptic	Gabapentin	5	5	0	2454	0	**1.00 (1.00–1.00)**	14	9	5	1912	1	0.75 (0.55–0.94)
Antidepressant	(Es-)citalopram	63	62	1	2388	8	**0.93 (0.89–0.98)**	27	24	3	1895	5	**0.86 (0.76–0.95)**
	Fluoxetine	16	16	0	2439	4	**0.89 (0.78–1.00)**	3	3	0	1924	0	**1.00 (1.00–1.00)**
	Venlafaxine	37	35	2	2420	2	**0.95 (0.89–1.00)**	12	12	0	1914	1	**0.96 (0.88–1.00)**
Antipsychotic	Quetiapine	13	13	0	2444	2	**0.93 (0.83–1.00)**	3	3	0	1924	0	**1.00 (1.00–1.00)**

Note: Kappa values in bold show an almost perfect agreement (κ > 0.8).

Abbreviations: CI, confidence interval; FN, false negative (metabolite detected but no self-reported use); FP, false positive (metabolite not detected but self-reported use); TN, true negative (metabolite not detected and no self-reported use); TP, true positive (metabolite detected and self-reported use).

TABLE A5 Concordance between self-reported drug use and measured metabolites with kappa (κ) values (95%CI) stratified for age groups

		< 65 years (n = 3216)						≥ 65 years (n = 1170)					
Drug class	**Drug**	**Regular users**	**TP**	**FP**	**TN**	**FN**	**Cohen's kappa (κ) (95% CI)**	**Regular users**	**TP**	**FP**	**TN**	**FN**	**Cohen's kappa (κ) (95% CI)**
Antihypertensive	Hydrochlorothiazide	139	73	66	3075	2	0.67 (0.60–0.75)	216	95	121	948	6	0.55 (0.48–0.61)
	Metoprolol	41	35	6	3164	11	0.80 (0.71–0.89)	60	57	3	1101	9	**0.90 (0.84–0.96)**
	Candesartan	113	106	7	3100	3	**0.95 (0.92–0.98)**	166	163	3	1003	1	**0.99 (0.97–1.00)**
	Olmesartan	14	11	3	3202	0	**0.88 (0.74–1.00)**	24	24	0	1146	0	**1.00 (1.00–1.00)**
	Valsartan	38	35	3	3177	1	**0.95 (0.89–1.00)**	60	59	1	1107	3	**0.97 (0.93–1.00)**
Uricostatic	Allopurinol	37	32	5	3165	14	0.77 (0.67–0.87)	55	53	2	1106	9	**0.90 (0.84–0.96)**
Antihyperglycemic	Metformin	41	33	8	3172	3	**0.86 (0.77–0.94)**	68	56	12	1102	0	**0.90 (0.84–0.96)**
	Sitagliptin	17	11	6	3199	0	0.78 (0.62–0.95)	29	20	9	1141	0	**0.81 (0.69–0.93)**
Antiepileptic	Gabapentin	8	4	4	3207	1	0.61 (0.30–0.93)	11	10	1	1159	0	**0.95 (0.86–1.00)**
Antidepressant	(Es-)citalopram	62	59	3	3143	11	**0.89 (0.84–0.95)**	28	27	1	1140	2	**0.95 (0.89–1.00)**
	Fluoxetine	14	14	0	3199	3	**0.90 (0.79–1.00)**	5	5	0	1164	1	**0.91 (0.73–1.00)**
	Venlafaxine	33	32	1	3181	2	**0.95 (0.90–1.00)**	16	15	1	1153	1	**0.94 (0.85–1.00)**
Antipsychotic	Quetiapine	14	14	0	3200	2	**0.93 (0.84–1.00)**	2	2	0	1168	0	**1.00 (1.00–1.00)**

Note: Kappa values in bold show an almost perfect agreement ($\kappa > 0.8$).

Abbreviations: CI, confidence interval; FN, false negative (metabolite detected but no self-reported use); FP, false positive (metabolite not detected but self-reported use); TN, true negative (metabolite not detected and no self-reported use); TP, true positive (metabolite detected and self-reported use).

3.2 Over- and undertreatment with levothyroxine- Findings of the population-based Rhineland Study

Original Article

Over- and Undertreatment With Levothyroxine

Findings of the Population-Based Rhineland Study

Nersi Alaeddin, Rutchanna M.S. Jongejan, Julia C. Stingl, Yolanda B. de Rijke, Robin P. Peeters, Monique M.B. Breteler, Folgerdiena M. de Vries

Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany: Nersi Alaeddin, Prof. Dr. Dr. Monique M.B. Breteler, Dr. Folgerdiena M. de Vries

Department of Clinical Chemistry, Erasmus MC University Medical Center, Rotterdam, The Netherlands: Dr. Rutchanna M.S. Jongejan

Department of Internal Medicine, Erasmus MC University Medical Center, Rotterdam, The Netherlands: Dr. Rutchanna M.S. Jongejan, Prof. Dr. Dr. Robin P. Peeters

Institute of Clinical Pharmacology, Faculty of Medicine, RWTH Aachen, Germany: Prof. Dr. med. Julia C. Stingl

Academic Centre for Thyroid Diseases, Erasmus MC University Medical Center, Rotterdam, The Netherlands: Prof. Dr. Yolanda B. de Rijke, Prof. Dr. Dr. Robin P. Peeters

Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Bonn, Germany: Prof. Dr. Dr. Monique M.B. Breteler

Summary

Background: Levothyroxine is a very commonly prescribed drug, and treatment with it is often insufficient or excessive. Nonetheless, there have been only a few reports on the determinants of inadequate levothyroxine treatment.

Methods: Data from 2938 participants in the population-based Rhineland Study were analyzed. Putative determinants of inadequate levothyroxine treatment (overtreatment, thyrotropin level <0.56 mU/L; undertreatment, thyrotropin level >4.27 mU/L) were studied with logistic regression. The determinants of the levothyroxine dose were assessed with linear regression.

Results: Overall, 23% of the participants (n = 662) stated that they were taking levothyroxine. Among these participants, 18% were overtreated and 4% were undertreated. Individuals over 70 years of age and above were four times as likely to be overtreated (OR = 4.05, 95% CI [1.20; 13.72]). Each rise in the levothyroxine dose by 25 µg was associated with an increased risk of overtreatment (OR = 1.02, 95% CI [1.02; 1.03]) and of undertreatment (OR = 1.02, 95% CI [1.00; 1.03]). Well-controlled participants (normal thyrotropin levels 0.56–4.27 mU/L) received a lower levothyroxine dose (1.04 ± 0.5 µg/kg/d) than overtreated (1.40 ±0.5 µg/kg/d) or undertreated (1.37 ±0.5 µg/kg/d) participants. No association was found between sociodemographic factors or comorbidities and the levothyroxine dose. Iodine supplementation was associated with a lower daily dose (β = −0.19, 95% CI [−0.28; −0.10]), while three years or more of levothyroxine exposure was associated with a higher daily dose (β = 0.24, 95% CI [0.07; 0.41]).

Conclusion: Levothyroxine intake was high in our sample, and suboptimal despite monitoring. Our findings underscore the need for careful dosing and for due consideration of deintensification of treatment where appropriate.

Cite this as:
Alaeddin N, Jongejan RMS, Stingl JC, de Rijke YB, Peeters RP, Breteler MMB, de Vries FM: Over- and undertreatment with levothyroxine—findings of the population-based Rhineland Study. Dtsch Arztebl Int 2023; 120: 711–8. DOI: 10.3238/arztebl.m2023.0192

Levothyroxine (LT4), the linchpin of thyroid hormone replacement therapy, is highly effective, inexpensive, and easy to administer (1, 2). LT4 use is increasing in many countries (3), most likely due to the increase in treatment of mild subclinical hypothyroidism (4, 5). In 2019, LT4 was the fourth most prescribed drug in Germany, with almost nine million prescriptions (6). A population-based study in Germany (age range 20 to > 80 years) reported a prevalence of LT4 use of 11%, while the Rhineland Study (age range 30–95 years) stated a prevalence of 24% (7–9). Studies in other European countries have reported prevalence rates of only 3–5% (10–12). These discrepancies may be attributable partly to regional differences in thyroid function parameters, thyroid diseases, or treatment protocols (13, 14).

The LT4 dosage is usually based on the serum level of thyrotropin (TSH). TSH must be monitored closely to avoid overtreatment, which causes high healthcare costs and adverse effects, or undertreatment, which has little clinical benefit (15, 16). Importantly, TSH levels outside the reference range are associated with adverse health outcomes, e.g., iatrogenic hyperthyroidism, increased cardiovascular morbidity/mortality, elevated fracture risk, and cognitive dysfunction (17–19). This is particularly true in older patients with suppressed TSH (20).

Despite the potential health risks, high rates of overtreatment (14–20%) and undertreatment (10–27%) have been described (10–12). However, these reports come from countries with a low prevalence of LT4 use compared with Germany. To date, only two studies have examined the quality of LT4 treatment in Germany. One of these (data from the period 1997–2001) reported over- and undertreatment rates of 19.5% and 10%, respectively (21), while the

TABLE 1

Characteristics of the study population

	Controlled	Overtreated	Undertreated	p*1	p*2
Participants, N (%)	518 (78.2)	117 (17.7)	27 (4.1)	< 0.001	< 0.001
Age (years), M (SD)	58.2 (13.9)	58.8 (13.6)	59.4 (15.5)	0.675	0.721
Sex (women), N (%)	435 (84.0)	96 (82.1)	21 (77.8)	0.659	0.432
Education, N (%)					
Low	16 (3.1)	2 (1.8)	0 (0.0)	0.474	–
Middle	276 (54.1)	57 (50.4)	17 (63.0)	Ref.	Ref.
High	218 (42.7)	54 (47.8)	10 (37.0)	0.369	0.457
Smoking, N (%)					
Never	216 (44.5)	50 (45.0)	11 (42.3)	Ref.	Ref.
Former	219 (45.2)	42 (37.8)	10 (38.5)	0.371	0.775
Current	50 (10.3)	19 (17.1)	5 (19.2)	0.116	0.261
BMI (kg/m^2)	26.5 (4.9)	25.7 (4.7)	25.9 (4.7)	0.064	0.439
TSH (mU/L), M (SD)	1.6 (0.8)	0.3 (0.2)	7.3 (3.4)	0.074	0.311
Diabetes, N (%)	29 (5.7)	7 (6.0)	2 (7.4)	0.961	0.787
Hypertension, N (%)	238 (47.1)	49 (42.2)	14 (51.9)	0.196	0.807
CVD, N (%)	54 (10.5)	10 (8.6)	5 (18.5)	0.425	0.274
CKD, N (%)	26 (5.4)	9 (8.0)	4 (14.8)	0.414	0.089
LT4 intake duration, N (%)					
≤ 12 months	29 (5.6)	8 (6.9)	3 (11.1)	Ref.	Ref.
13–36 months	80 (15.4)	16 (13.8)	2 (7.4)	0.481	0.277
> 36 months	409 (79.0)	92 (79.3)	22 (81.5)	0.584	0.277
Iodine supplementation, N (%)	145 (28.5)	31 (26.5)	3 (11.1)	0.610	0.052
Polypharmacy, N (%)	161 (31.1)	38 (32.5)	9 (33.3)	0.907	0.956
Global cognition (z-score), M (SD)	−0.1 (0.6)	−0.1 (0.7)	−0.1 (0.8)	0.732	0.641

Treatment status "controlled" (TSH 0.56–4.27 mU/L), "overtreated" (TSH < 0.56 mU/L), "undertreated" (TSH > 4.27 mU/L). Group differences were calculated with logistic regression adjusted for age and sex (age and sex were only adjusted for the other respectively).

*1 Adjusted for age and sex (overtreated compared with controlled)

*2 Adjusted for age and sex (undertreated compared with controlled)

BMI, Body mass index; CKD, chronic kidney disease; CVD, cardiovascular disease; LT4, levothyroxine; M, mean; N, number of participants; Ref, reference group; SD, standard deviation; TSH: thyrotropin.

other (data from the years 2005–2018) reported only the cumulative risk (overtreatment 1.3%, undertreatment 3%) (22). The prevalence, however, was not reported, so the current burden in Germany remains unclear (22). A German study published in 2020 found that TSH levels are poorly monitored in LT4 users. Investigation of the current extent of over- and undertreatment is therefore needed (8).

Evidence on the determinants of over- and undertreatment and LT4 dose is also limited. Longer LT4 exposure duration and higher LT4 dose were associated with overtreatment, while men and younger persons were more likely to be undertreated (12). Age, sex, and body weight were associated with LT4 dosage, but these studies were conducted in older, obese patients or in patients who had undergone thyroidectomy (23–25).

The aim of the study described herein was to investigate the prevalence and determinants of LT4 over- and undertreatment, together with the determinants of LT4 dose, in a large-scale population-based study. Furthermore we evaluated information on the initiation, duration, and monitoring of treatment among LT4 users.

Methods

Study population

We used data from the Rhineland Study, a community-based cohort *(eMethods, eTable 1)*. All residents (≥ 30 years) of two geographically defined areas in Bonn, Germany were invited to take part. The sole inclusion criterion was possession of sufficient German language skills to provide informed consent. The baseline data of the first 3000 participants (March 2016

to February 2020) with measured serum TSH were used. We excluded 62 participants due to incomplete TSH measurements (n = 2), missing medication data (n = 54), or because they were taking drugs that affected thyroid hormone levels (amiodarone/lithium; n = 6), so 2938 persons were included in the analyses. We also conducted a brief online survey in 2022 to obtain additional information on regular LT4 users *(eMethods, eTable 2)*.

TSH assessment

Blood samples were taken in the morning after a 10-hour fast. The laboratory defined the reference range of TSH as 0.56–4.27 mU/L (26). Details of blood collection and TSH measurement/reference range can be found in the *eMethods*.

LT4 treatment

All participants were asked to bring the original packaging of all medications they were currently using and had taken as needed in the past year. Data were collected by interview, documenting name, dosage, and current prescription status (9, 27). LT4 treatment status was categorized by TSH levels: adequate, i.e., controlled (0.56–4.27 mU/L), or inadequate, i.e. overtreatment (< 0.56 mU/L) or undertreatment (> 4.27 mU/L).

Statistical analysis

The participants' characteristics were summarized using descriptive statistics. Group differences were calculated using logistic regression (adjusted for age and sex). Multinomial logistic regression was performed to identify possible determinants *(eTable 1)* of over- and undertreatment in LT4 users (reference group: controlled participants) in a fully adjusted model. Multivariable linear regression was then used to identify predictors of LT4 dose (µg/kg/d) in a fully adjusted model. Statistical analyses were conducted using RStudio (version 4.1.1).

Results

Study population

The participants' characteristics are presented in *Table 1 and eTable 2*. The persons included (n = 2938) were on average 55 ± 14.4 years old (range 30–95; 56.5% women) and did not differ significantly from those who were excluded (n = 62) in terms of age (57 ± 14.9 years, range 30–87; p = 0.187) or sex ratio (women n = 34, 54.8%; p = 0.417).

Overtreatment and undertreatment with LT4

Regular LT4 use was reported by 22.5% of the participants. Users were older than non-users (58.3 vs. 54.1 years; p < 0.001), and prevalence was higher in women than in men (33.3% vs. 8.6%; p < 0.001). Among LT4 users (n = 662), 78.2% were controlled (n = 518), while 21.8% (n = 144) were inadequately treated, of whom 17.7% (n = 117) were overtreated and 4.1% (n = 27) were undertreated.

Logistic regression showed that persons aged ≥ 70 years were four times more likely to be overtreated (odds ratio 4.05; 95% confidence interval [1.20; 13.72]) than those who were younger, and that increasing the LT4 dose by 25 µg/d increased the likelihood of both overtreatment (OR 1.02; [1.02; 1.03]) and undertreatment (OR 1.02; [1.00; 1.03]) *(Table 2)*.

LT4 dose

Controlled persons had lower daily doses of LT4 (1.04 ± 0.5 µg/kg/d) than overtreated (1.40 ± 0.5 µg/kg/d) and undertreated (1.37 ± 0.5 µg/kg/d) persons, but adjustment for age and sex revealed no significant differences *(Table 1)*. *Figure 1* illustrates the LT4 doses (µg/kg/d) in relation to the TSH levels (mU/L) and shows individuals with very high and very low doses in all treatment groups. We found no association of sociodemographic factors or comorbidities with the daily LT4 dose. Iodine supplementation (β = –0.19; [–0.28; –0.10], p < 0.001) was associated with lower LT4 dose, and LT4 exposure duration (β= 0.24; [0.07; 0.41), p = 0.001) of ≥ 3 years was associated with a higher dose *(Table 3)*.

Online survey

The results of the survey are shown in *eTable 3*. The LT4 users who were included (n = 456; mean age 56.0 ± 13.0 years, range 30–94; 83.1% women) did not differ from those who were excluded (n = 206; mean age 56.0 ± 12.4 years, range 31–88; 84.4% women) in terms of age (p = 0.913) or sex (p = 0.613). Participants were predominantly long-term users (21.3 ± 12.2 years) and 60.4% reported having their TSH levels monitored every 6–12 months.

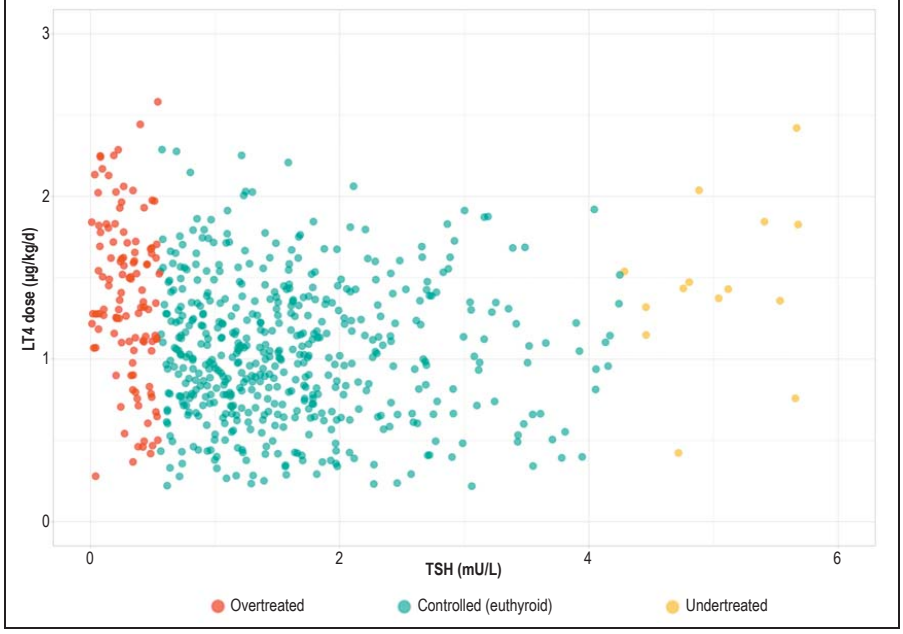
Discussion

We investigated the prevalence and determinants of overtreatment, undertreatment, and LT4 dosage in a large population-based cohort. A high proportion of participants, mainly women (women 33%; men 9%), reported taking LT4 (23%). Of these, 18% were overtreated and 4% undertreated. Older age was associated with overtreatment, while higher LT4 dose was associated with both overtreatment and undertreatment. Iodine supplementation was associated with lower LT4 dosage, whereas longer LT4 intake (≥ 3 years) was associated with higher doses.

LT4 is the most commonly used drug in our cohort (9). The frequency of use was higher in women than in men and increased with age, which was to be expected based on the prevalence of thyroid disease in these groups (11, 17, 28, 29). Importantly, the adverse health consequences of overtreatment are most pronounced in the elderly (20).

An increase in LT4 prescriptions has been observed worldwide, apparently mainly due to increased treatment of subclinical hypothyroidism with mild TSH elevation (30). Although treating mildly elevated TSH levels (< 10 mIU/L) is not recommended in the current guidelines (31), an American study found a

FIGURE



LT4 dose (µg/kg/d) compared with TSH (mU/L), stratified by treatment status
LT4, Levothyroxine; TSH, thyrotropin

median TSH of 5.3 mIU/L in 9331 patients newly started on LT4 treatment (5). This is worrying and shows how remarkable it is that the prevalence of LT4 use in our study is seven times higher than in other European studies (3.1–4.4%) (10–12) and more than twice as high as in other German population-based studies (~11%) (7, 8). Possible reasons for the differences between our study and other German studies are the period of data collection (2000–2016), different age and sex distributions, and regional variations in prescribing patterns, iodine availability, or thyroid disease (7, 8, 13, 14).

One explanation for the overall high prevalence of LT4 use in Germany may be the frequent use of TSH measurement and thyroid ultrasound (8). In Germany, there appears to be a strong focus on the thyroid both in detection of morphological changes and in drug therapy.

Indeed, the annual rate of thyroid surgery (109/100 000) is high compared with England (27/100 000) or the Netherlands (16/100 000), and thyroid hormone prescriptions increased by 40% between 2010 and 2019, when almost 9 million prescriptions were issued (6, 32). Additionally, a case-based survey found that German general practitioners were more likely to prescribe LT4 for patients with subclinical hypothyroidism than their colleagues in other European countries (33).

Approximately 18% of LT4 users were overtreated and 4% undertreated. Although the prevalence of LT4 use in our study is higher than in other studies, our results are comparable regarding overtreatment (14–20%), though not for undertreatment (10–27%) (10–12, 21). Given the high use of LT4 in our population, we expected higher rates of overtreatment. Perhaps we underestimate the prevalence of overtreatment *(Figure 1)*, because controlled persons with low LT4 doses and low TSH levels could be overtreated as TSH levels would probably remain within the reference range after LT4 discontinuation. Whether these individuals require treatment cannot be conclusively established on the basis of our data. In some cases, e.g., patients with thyroid cancer (34), very low TSH is desirable so the levels are deliberately kept low. However, no individuals in our sample self-reported thyroid cancer, so the high prevalence of overtreatment cannot be justified in this way.

Similar to another German study (8), almost 60% of LT4 users in the Rhineland Study reported having their TSH levels monitored every 6–12 months, with no noticeable difference between controlled and inadequately treated persons *(eTable 3)*. This demonstrates that frequent monitoring does not necessarily prevent inadequate treatment.

TABLE 2

Determinants of LT4 overtreatment and undertreatment (n = 557)

Status	Determinant	OR	[95% CI]	p
Overtreated	Age 40–49 years (vs. 30–39 years)	0.89	[0.34; 2.35]	0.810
	Age 50–59 years (vs. 30–39 years)	1.87	[0.73; 4.75]	0.189
	Age 60–69 years (vs. 30–39 years)	2.73	[0.96; 7.74]	0.060
	Age ≥ 70 years (vs. 30–39 years)	4.05	1.20; 13.72]	0.025
	Sex (men vs. women)	0.84	[0.43; 1.65]	0.616
	Education (low vs. middle)	0.29	[0.03; 2.52]	0.262
	Education (high vs. middle)	1.22	[0.74; 2.00]	0.438
	Smoking (former vs. never)	0.61	[0.36; 1.04]	0.068
	Smoking (current vs. never)	1.42	[0.69; 2.92]	0.337
	BMI (kg/m^2, increase per unit)	0.95	[0.90; 1.01]	0.077
	Diabetes (yes vs. no)	0.63	[0.21; 1.94]	0.420
	Hypertension (yes vs. no)	0.79	[0.44; 1.41]	0.422
	CVD (yes vs. no)	0.51	[0.19; 1.37]	0.182
	CKD (yes vs. no)	1.50	[0.53; 4.24]	0.444
	Iodine supplementation (yes vs. no)	1.11	[0.63; 1.96]	0.725
	Polypharmacy (yes vs. no)	1.07	[0.60; 1.92]	0.812
	LT4 dose (per 25 µg increase)	1.02	[1.02; 1.03]	< 0.001
	LT4 intake 13–36 months (vs. 0–12 months)	0.59	[0.18; 1.96]	0.389
	LT4 intake > 36 months (vs. 0–12 months)	0.50	[0.18; 1.42]	0.192
	Global cognition (z-score, per SD)	1.23	[0.72; 2.09]	0.445
Undertreated	Age 40–49 years (vs. 30–39 years)	0.64	[0.12; 3.53]	0.607
	Age 50–59 years (vs. 30–39 years)	1.58	[0.34; 7.41]	0.560
	Age 60–69 years (vs. 30–39 years)	1.42	[0.23; 8.76]	0.703
	Age ≥ 70 years (vs. 30–39 years)	1.72	[0.21; 14.27]	0.617
	Sex (men vs. women)	1.42	[0.47; 4.32]	0.539
	Education (low vs. middle)	–	–	–
	Education (high vs. middle)	0.85	[0.35; 2.09]	0.725
	Smoking (former vs. never)	0.78	[0.29; 2.09]	0.623
	Smoking (current vs. never)	1.93	[0.56; 6.64]	0.299
	BMI (kg/m^2, increase per unit)	0.96	[0.87; 1.06]	0.418
	Diabetes (yes vs. no)	1.16	[0.20; 6.63]	0.867
	Hypertension (yes vs. no)	0.97	[0.33; 2.81]	0.951
	CVD (yes vs. no)	3.04	[0.80; 11.56]	0.103
	CKD (yes vs. no)	2.08	[0.42; 10.39]	0.370
	Iodine supplementation (yes vs. no)	0.27	[0.06; 1.27]	0.097
	Polypharmacy (yes vs. no)	0.67	[0.21; 2.13]	0.494
	LT4 dose (per 25 µg increase)	1.02	[1.00; 1.03]	0.017
	LT4 intake 13–36 months (vs. 0–12 months)	0.26	[0.03; 2.23]	0.220
	LT4 intake > 36 months (vs. 0–12 months)	0.42	[0.08; 2.23]	0.312
	Global cognition (z-score, per SD)	1.37	[0.55; 3.46]	0.499

The sample size is based on persons with complete data on all determinants.
BMI, Body mass index; CI, confidence interval; CKD, chronic kidney disease; CVD, cardiovascular disease; LT4, levothyroxine; n, number of participants; OR, odds ratio; SD, standard deviation; vs., versus

TABLE 3

Determinants of LT4 dose (increase per unit µg/kg/d), n = 556

Determinant	β	[95% CI]	p
Age 40–49 years (vs. 30–39 years)	0.06	[−0.09; 0.22]	0.437
Age 50–59 years (vs. 30–39 years)	0.06	[−0.09; 0.21]	0.433
Age 60–69 years (vs. 30–39 years)	−0.09	[−0.25; 0.07]	0.267
Age ≥ 70 years (vs. 30–39 years)	−0.09	[−0.26; 0.08]	0.297
Sex (men vs. women)	−0.02	[−0.13; 0.10]	0.783
Education (low vs. middle)	−0.04	[−0.30; 0.22]	0.744
Education (high vs. middle)	−0.00	[−0.09; 0.08]	0.941
Smoking (former vs. never)	0.05	[−0.04; 0.13]	0.285
Smoking (current vs. never)	0.11	[−0.02; 0.24]	0.106
Diabetes (yes vs. no)	0.07	[−0.11; 0.26]	0.431
Hypertension (yes vs. no)	−0.06	[−0.15; 0.03]	0.209
CVD (yes vs. no)	0.03	[−0.12; 0.17]	0.720
CKD (yes vs. no)	0.08	[−0.10; 0.27]	0.371
Iodine supplementation (yes vs. no)	−0.19	[−0.28; −0.10]	< 0.001
Polypharmacy (yes vs. no)	0.03	[−0.06; 0.13]	0.493
TSH (mU/L, increase per unit)	−0.00	[−0.03; 0.02]	0.704
LT4 intake duration 13–36 months (vs. 0–12 months)	0.02	[−0.17; 0.21]	0.838
LT4 intake duration > 36 months (vs. 0–12 months)	0.24	[0.07; 0.41]	0.006

The sample size is based on individuals with complete data on all determinants.
BMI, Body mass index; CKD, chronic kidney disease; CVD, cardiovascular disease; LT4, levothyroxine; n, number of participants; TSH, thyrotropin; vs., versus

The likelihood of overtreatment was high in individuals ≥ 70 years, which is unsurprising as it has been reported that overtreatment is common in the elderly (35). Complex LT4 treatment regimens, with varying dosages across weekdays to achieve optimal titration, can become challenging with increasing age, especially as non-adherence rises with age (17). Among the LT4 users with suppressed TSH levels, 27% were aged ≥ 70 years. Importantly, suppressed TSH is particularly strongly associated with adverse health outcomes in the elderly (20). This makes the finding that 23% of all participants used LT4 all the more important.

Reducing the number of unnecessary LT4 prescriptions may improve health status and reduce healthcare costs. Future studies should aim to understand what factors contribute to use of LT4 by this extremely high proportion of people. In agreement with another study (12), the probability of over- and undertreatment rose with increasing LT4 dose. One can only speculate about the possible reasons for undertreatment despite high dosage. One possible explanation is lack of adherence to treatment, or reluctance on the part of physicians to increase the dose beyond a certain point for fear of adverse events. In contrast to a previous study, we did not find that men were more often undertreated than women, but we did observe a trend in that direction (12). This could be because thyroid disease is more common in women, and women more frequently receive TSH tests (8).

Both overtreated and undertreated participants had higher mean daily doses than controlled users. No associations were found between either sociodemographic factors or comorbidities and LT4 dosage. However, iodine supplementation was associated with lower daily doses, while LT4 exposure duration of ≥ 3 years was associated with higher daily doses.

Although there was no significant association between age and LT4 dose, younger persons tended to receive higher doses and older persons to receive lower doses, which is consistent with recent evidence that older persons should start with a low dose (24, 36).

Iodine, an important micronutrient, is known to control thyroid function by reducing the thyroid gland's response to TSH. In high concentrations, iodine inhibits thyroid hormone secretion. Especially in persons with pre-existing thyroid disease, iodine can induce hypo- or hyperthyroidism (37). Therefore, correct dose adjustment in iodine supplementation is all the more important.

One possible reason for the association between the duration of LT4 intake and higher dosage is that treatment for thyroid hormone deficiency is usually started at a low dose and then increased by dose titration to achieve target TSH levels. Alternatively, thyroid function may decline progressively in patients who initially have subclinical hypothyroidism.

Strengths and limitations

One of the strengths of our study is the examination of both overtreatment and undertreatment in a large population-based cohort. Our extensive data allowed us to analyze various determinants, and self-reported medication data may better reflect actual use than secondary data. Although self-reported medication data may introduce reporting bias, we validated the reliability of our data (27).

Potential limitations include the fact that treatment adherence could not be considered. Furthermore, as is often the case in epidemiological studies (21), only one TSH measurement time point was available, so that our results cannot account for any TSH fluctuations (38). We did not have detailed information on whether and when dose adjustments were made. According to participants' reports, however, the last dose adjustment had taken place on average 6 years earlier. Moreover, we do not have longitudinal data, so we could not follow changes in TSH levels or general health. Finally, our population may be "healthier," which would limit the generalizability of our results. However, the prevalences of hypertension and polypharmacy and the age and sex distributions are all comparable with the German population (9, 27).

The prevalence of LT4 use in our population was very high and was suboptimal in almost a quarter of the participants despite frequent TSH monitoring. This high proportion of LT4 use is probably due to overtreatment in the vast majority of participants and, assuming that 18% of participants have suppressed TSH, will contribute to adverse health outcomes.

Conclusion

Our report suggests that the focus should be not only on intensification of treatment, but also on deintensification. Furthermore, the strategy for monitoring should be reconsidered, as it does not appear to lead to high-quality care at present.

Study funding
This work was supported by a grant provided by the Federal Institute for Drugs and Medical Devices (BfArM), Bonn, Germany (V-15646/68502/2014–2016 and V-17746/68502/2018–2020).

Data sharing
The datasets for this manuscript are not publicly available because of data protection regulations. Access to data can, however, be provided to scientists in accordance with the Rhineland Study's Data Use and Access Policy. Requests to access the datasets should be directed to Prof. Dr. Dr. Monique M.B. Breteler, RS-DUAC@dzne.de.

Conflict of interest statement
The authors declare that no conflict of interest exists.

Manuscript received on 1 March 2023, revised version accepted on 7 August 2023.

References

1. Chaker L, Bianco AC, Jonklaas J et al. Hypothyroidism. Lancet 2017;390:1550–62.
2. Eligar V, Taylor PN, Okosieme OE et al. Thyroxine replacement: a clinical endocrinologist's viewpoint. Ann Clin Biochem 2016;53:421–33.
3. Rodriguez-Gutierrez R, Maraka S, Ospina NS et al. Levothyroxine overuse: time for an about face? Lancet Diabetes Endocrinol 2017;5:246–8.
4. Taylor PN, Iqbal A, Minassian C et al. Falling Threshold for Treatment of Borderline Elevated Thyrotropin Levels—Balancing Benefits and Risks. JAMA Intern Med 2014;174:32.
5. Brito JP, Ross JS, El Kawkgi OM et al. Levothyroxine Use in the United States, 2008–2018. JAMA Intern Med 2021;181:1402–5.
6. Schwabe U, Wolf-Dieter Ludwig. Arzneiverordnungs-Report 2020. Springer Berlin Heidelberg, 2020, 825–31.
7. Khattak RM, Ittermann T, Nauck M et al. Monitoring the prevalence of thyroid disorders in the adult population of Northeast Germany. Popul Health Metr 2016;14, DOI: 10.1186/s12963–016–0111–3.
8. Kiel S, Ittermann T, Völzke H et al. Frequency of thyroid function tests and examinations in participants of a population-based study. BMC Health Serv Res 2020;20:70.
9. de Vries FM, Stingl JC, Breteler MMB. Polypharmacy, potentially inappropriate medication and pharmacogenomics drug exposure in the Rhineland Study. Br J Clin Pharmacol 2021;87:2732–56.
10. Wouters HJCM, Slagter SN, Muller Kobold AC et al. Epidemiology of thyroid disorders in the Lifelines Cohort Study (the Netherlands). Naitza S (ed.). PLoS One 2020;15:e0242795.
11. Janett-Pellegri C, Wildisen L, Feller M et al. Prevalence and factors associated with chronic use of levothyroxine: A cohort study. Lupattelli A (ed.). PLoS One 2021;16:e0261160.
12. Okosieme OE, Belludi G, Spittle K et al. Adequacy of thyroid hormone replacement in a general population. Qjm 2011;104:395–401.
13. Girschik C, Muchalla P, Kowall B et al. Regionale Unterschiede von Schilddrüsenfunktionsparametern: ein Vergleich europäischer Kohortenstudien. Das Gesundheitswes 2022:1–2.
14. Meisinger C, Ittermann T, Wallaschofski H et al. Geographic variations in the frequency of thyroid disorders and thyroid peroxidase antibodies in persons without former thyroid disease within Germany. Eur J Endocrinol 2012;167:363–71.
15. Lyu H, Xu T, Brotman D et al. Overtreatment in the United States. PLoS One 2017;12:1–11.
16. Kearney M, Treadwell J, Marshall M. Overtreatment and undertreatment: time to challenge our thinking. Br J Gen Pract 2017;67:442–3.
17. Effraimidis G, Watt T, Feldt-Rasmussen U. Levothyroxine Therapy in Elderly Patients With Hypothyroidism. Front Endocrinol (Lausanne) 2021;12:1–12.
18. Blum MR, Bauer DC, Collet T-H et al. Subclinical Thyroid Dysfunction and Fracture Risk. JAMA 2015;313:2055.
19. Collet T-H, Gussekloo J, Bauer DC et al. Subclinical Hyperthyroidism and the Risk of Coronary Heart Disease and Mortality. Arch Intern Med 2012;172:1–7.
20. Biondi B, Cooper DS. Subclinical Hyperthyroidism. N Engl J Med 2018;378:2411–9.
21. Hannemann A, Friedrich N, Haring R et al. Thyroid function tests in patients taking thyroid medication in Germany: Results from the population-based Study of Health in Pomerania (SHIP). BMC Res Notes 2010;3:227.
22. Kostev K. Frequency of over- and under-treatment with levothyroxine in primary care in Germany. Eur J Endocrinol 2022;186:L5.
23. Mele C, Tagliaferri MA, Pagano L et al. Levothyroxine Replacement in Obese Adults: The Role of Metabolic Variables and Aging on Thyroid Testing Abnormalities. J Clin Endocrinol Metab 2019;104:6265–74.
24. Jonklaas J. Sex and Age Differences in Levothyroxine Dosage Requirement. Endocr Pract 2010;16:71–9.
25. Devdhar M, Drooger R, Pehlivanova M et al. Levothyroxine Replacement Doses Are Affected by Gender and Weight, But Not Age. Thyroid 2011;21:821–7.
26. Pellikaan K, Snijders F, Rosenberg AGW et al. Thyroid Function in Adults with Prader–Willi Syndrome; a Cohort Study and Literature Review. J Clin Med 2021;10:3804.
27. Alaeddin N, Stingl JC, Breteler MMB et al. Validation of self‐reported medication use applying untargeted mass spectrometry‐based metabolomics techniques in the Rhineland study. Br J Clin Pharmacol 2022;88:2380–95.
28. Canaris GJ, Manowitz NR, Mayor G et al. The colorado thyroid disease prevalence study. Arch Intern Med 2000;160:526–34.
29. Åsvold BO, Vatten LJ, Bjøro T. Changes in the prevalence of hypothyroidism: The HUNT study in Norway. Eur J Endocrinol 2013;169:613–20.
30. Bekkering GE, Agoritsas T, Lytvyn L et al. Thyroid hormones treatment for subclinical hypothyroidism: a clinical practice guideline. BMJ 2019;365:l2006.
31. Peeters RP. Subclinical Hypothyroidism. Solomon CG (ed.). N Engl J Med 2017;376:2556–65.
32. Verburg FA. Is thyroid surgery performed too often in Germany? Nuklearmedizin 2015;54:101–5.
33. den Elzen WPJ, Lefèbre-van de Fliert AA, Virgini V et al. International variation in GP treatment strategies for subclinical hypothyroidism in older adults: a case-based survey. Br J Gen Pract 2015;65:e121–32.
34. Haugen BR, Alexander EK, Bible KC et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;26:1–133.
35. Somwaru LL, Arnold AM, Joshi N et al. High Frequency of and Factors Associated with Thyroid Hormone Over-Replacement and Under-Replacement in Men and Women Aged 65 and Over. J Clin Endocrinol Metab 2009;94:1342–5.
36. Gottwald-Hostalek U, Razvi S. Getting the levothyroxine (LT4) dose right for adults with hypothyroidism: opportunities and challenges in the use of modern LT4 preparations. Curr Med Res Opin 2022;0:1–6.
37. Leung AM, Braverman LE. Iodine-induced thyroid dysfunction. Curr Opin Endocrinol Diabetes Obes 2012;19:414–9.
38. Weeke J, Gundersen HJ. Circadian and 30 minutes variations in serum TSH and thyroid hormones in normal subjects. Acta Endocrinol (Copenh) 1978;89:659–72.

39. Schübel J, Voigt K, Uebel T. Erhöhter TSH-Wert in der Hausarztpraxis. S2k-Leitlinie. DEGAM Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin e.V. 2023 (accessed 04 August 2023)

Corresponding author
Dr. Folgerdiena M. de Vries
Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE)
Venusberg-Campus 1, Gebäude 99
53127 Bonn, Germany
Dianna.DeVries@dzne.de

Cite this as:
Alaeddin N, Jongejan RMS, Stingl JC, de Rijke YB, Peeters RP, Breteler MMB, de Vries FM: Over- and undertreatment with levothyroxine—findings of the population-based Rhineland Study. Dtsch Arztebl Int 2023; 120: 711–8. DOI: 10.3238/arztebl.m2023.0192

► **Supplementary material**
eMethods, eTables:
www.aerzteblatt.de/m2023.0192

Supplementary material to:

Over- and Undertreatment With Levothyroxine

Findings of the Population-Based Rhineland Study

by Nersi Alaeddin, Rutchanna M.S. Jongejan, Julia C. Stingl, Yolanda B. de Rijke, Robin P. Peeters, Monique M.B. Breteler, and Folgerdiena M. de Vries

Dtsch Arztebl Int 2023; 120: 711–8. DOI: 10.3238/arztebl.m2023.0192

eMETHODS

Study design

The Rhineland Study is an ongoing community-based, prospective cohort study. The participants are residents of two geographically defined areas in Bonn, Germany. Recruitment began in 2016. All residents aged ≥ 30 years were invited using contact information provided by the municipality. Participation was by invitation only, and invitations were issued regardless of the health status of those invited. The sole exclusion criterion was insufficient command of the German language to provide written informed consent. The study of (neurodegenerative) diseases and the identification of determinants and biomarkers of healthy aging is a primary objective of the Rhineland Study. Therefore, all participants underwent a standardized 8-hour in-depth phenotyping process, including cardiovascular health assessment, brain imaging, cognitive testing, metabolite profiling, and documentation of medication use. The data were collected through questionnaires, interviews, and the collection of various biomaterials such as blood, stool, urine, and hair samples. Approval to conduct the study was granted by the ethics committee of the Medical Faculty of the University of Bonn. The study protocols were conducted in accordance with the recommendations of the International Council for Harmonisation and the Good Clinical Practice standards. Written informed consent was obtained in accordance with the tenets of the Declaration of Helsinki. No financial incentives were offered to the participants.

Online survey

In addition to the data collected at baseline, we wanted to acquire more information about thyroid disease and thyroid hormone replacement therapy. Therefore, we initiated a short online survey (data collection: September 2022–March 2023). We asked all LT4 users (n = 662) to complete a questionnaire to obtain further information about the initiation, cause, duration, and monitoring of treatment and about the thyroid examinations performed. The questionnaire was completed by 456 of the 662 regular LT4 users.

TSH assessment

Venous blood was collected from participants who had fasted for at least 10 hours. The blood was transferred to S-Monovette tubes (7.5 mL) containing coagulation factor and incubated for 30 minutes for coagulation (room temperature). The tubes were centrifuged at 2000 × g and 4 °C for 15 minutes. The samples were then aliquoted (500 µL each) and transferred into 0.7-ml FluidX tubes. After aliquoting, all samples were immediately frozen at −80 °C. The TSH level in the serum samples was then measured using the Lumipulse G1200 (FujiRebio Inc., Ghent,

Belgium), a non-competitive chemiluminescent enzyme immunoassay (Erasmus MC, University Medical Center, Rotterdam, The Netherlands) . The TSH reference values were set by the laboratory at 0.56–4.27 mU/L. It should be noted that the measurement of TSH is instrument- and laboratory-dependent and therefore the reference values also depend on the methods, reagents, and calibration standards used. The TSH reference ranges set by laboratories therefore vary both internationally and within Germany, as noted in the current German guideline *Erhöhter TSH-Wert in der Hausarztpraxis* (Elevated TSH Levels in Primary Care) (39).

eTABLE 1

Definition of demographic and clinical characteristics

	Characteristic	Missing	Definition
General characteristics	Age group	0.0%	Age range 30–95 years: 30–39, 40–49, 50–59, 60–69, ≥ 70 years
	Sex	0.0%	Women, men
	Education	1.0%	Based on the International Standard Classification of Education 2011 (ISCED): low (lower secondary education or below), middle (upper secondary education to undergraduate university level), high (postgraduate university study)
	Smoking	5.9%	Persons who have never smoked, formerly smoked, or currently smoke
	Body mass index	0.4%	Body mass divided by square of body height (kg/m^2)
Comorbidities	Diabetes	0.9%	Self-reported physician diagnosis and/or glycated hemoglobin (HbA1c) (no diabetes < 6.5%; diabetes ≥ 6.5%), fasting glucose (no diabetes < 126 mg/dL; diabetes ≥ 126 mg/dL) measured in fasting morning blood, and/or intake of antidiabetics
	Hypertension	1.6%	Based on the 2018 European Society of Cardiology guidelines for the management of arterial hypertension: mean systolic blood pressure ≥ 140 mmHg and/or mean diastolic blood pressure ≥ 90 mmHg and/or antihypertensive drug use, irrespective of blood pressure
	Cardiovascular disease	0.4%	Based on a self-reported physician diagnosis of one or more of the following conditions: myocardial infarction, coronary artery disease, cardiac insufficiency, cardiac pacemaker, peripheral artery occlusive disease, stroke, surgery on large vessels such as aorta, carotid, or peripheral vessels
	Chronic kidney disease	5.5%	Estimated glomerular filtration rate based on cystatin C (no CKD ≥ 60 mL/min/1.73 m^2; CKD < 60 mL/min/1.73 m^2)
Medication	LT4 dosage (µg/kg/d)	0.4%	Daily dose of LT4 consumed, expressed in relation to body weight
	LT4 intake duration	0.0%	0–12 months, 13–36 months, > 36 months
	Iodine supplementation	0.4%	Regular intake of iodine (ATC H03CA01)
	Polypharmacy	0.0%	Regular use of ≥ 5 prescribed drugs
Cognition	Global cognition (z-standardized)	1.9%	Derived from a cognitive test battery assessing episodic verbal memory, working memory, executive function and processing speed

CKD, Chronic kidney disease; LT4, levothyroxine

eTABLE 2

Prevalence of self-reported thyroid disease ever diagnosed by a doctor

	All	LT4 treatment status				
		Controlled	Overtreated	Undertreated	p*1	p*2
Hypothyroidism, N (%)	310 (11.1)	190 (40.7)	36 (35.0)	7 (29.2)	0.264	0.270
Hyperthyroidism, N (%)	141 (5.1)	58 (12.4)	11 (10.7)	6 (25.0)	0.633	0.084
Hashimoto, N (%)	182 (6.5)	129 (27.6)	35 (34.0)	8 (33.3)	0.188	0.505
Basedow, N (%)	29 (1.0)	17 (3.6)	3 (2.9)	1 (4.2)	0.725	0.896
Goiter, N (%)	90 (3.2)	41 (8.8)	11 (10.7)	2 (8.3)	0.529	0.929

Group differences were calculated with logistic regression, adjusted for age and sex (age and sex were only adjusted for the other, respectively)
Treatment status controlled: TSH 0.56–4.27 mU/L; overtreated: TSH < 0.56 mU/L; undertreated: TSH > 4.27 mU/L
*1 Adjusted for age and sex (overtreated compared with controlled)
*2 Adjusted for age and sex (undertreated compared with controlled)
LT4, Levothyroxine; N, number of participants; TSH: thyrotropin

eTABLE 3

Results of the online survey (n=456)

	LT4 users	Missing	LT4 treatment status		
			Controlled	Overtreated	Undertreated
Participants, N	456		361	82	13
Sex, N (%)		0.0%			
Women	379 (83.1)		300 (83.1)	66 (80.5)	13 (100.0)
Men	76 (16.7)		60 (16.6)	16 (19.5)	0 (0.0)
Diverse	1 (0.2)		1 (0.3)	0 (0.0)	0 (0.0)
LT4 intake (years), M (SD)	21.3 (12.2)	9.2%	20.6 (12.0)	24.1 (12.9)	22.8 (9.9)
Diagnosis-based initiation of LT4, N (%)	450	1.2%			
Hypothyroidism	174 (38.7)		139 (39.1)	30 (36.6)	5 (38.5)
Benign struma	82 (18.2)		59 (16.6)	22 (26.8)	1 (7.7)
Hashimoto	122 (27.1)		94 (26.5)	23 (28.0)	5 (38.5)
Other diagnosis	51 (11.3)		44 (12.4)	5 (6.1)	2 (15.4)
Unknown	21 (4.7)		19 (5.4)	2 (2.4)	0 (0.0)
TSH monitoring frequency, N (%)		19.3%			
Every 6 months	61 (16.6)		43 (14.6)	13 (21.0)	5 (45.5)
Yearly	161 (43.8)		131 (44.4)	25 (40.3)	5 (45.5)
Every 1–2 years	77 (20.9)		63 (21.4)	14 (22.6)	0 (0.0)
Irregularly	56 (15.2)		48 (16.3)	7 (11.3)	1 (9.1)
No monitoring	13 (3.5)		10 (3.4)	3 (4.8)	0 (0.0)
Most recent LT4 dose adjustment in years, M (SD)	6.2 (6.0)	24.1%	6.5 (6.1)	5.6 (6.1)	3.1 (1.7)
Thyroid examinations performed		9.0%			
Biopsy	54 (11.8)		42 (11.6)	11 (13.4)	1 (7.7)
Ultrasound	389 (85.3)		305 (84.5)	72 (87.8)	12 (92.3)
Scintigraphy	275 (60.3)		218 (60.4)	50 (61.0)	7 (53.8)

Treatment status controlled: TSH 0.56–4.27 mU/L; overtreated: TSH < 0.56 mU/L; undertreated: TSH > 4.27 mU/L
LT4, Levothyroxine; M, mean; N, number of participants; SD, standard deviation; TSH, thyrotropin

3.3 Prevalence and determinants of over- and undertreatment among users of antihypertensive drugs in the general population: the Rhineland Study

ESC European Society of Cardiology
European Journal of Preventive Cardiology (2023) **00**, 1–4
https://doi.org/10.1093/eurjpc/zwad274

RESEARCH LETTER
Population science

Prevalence and determinants of over- and undertreatment among users of antihypertensive drugs in the general population: the Rhineland Study

Nersi Alaeddin [1], Gökhan Pehlivan [1], Julia C. Stingl [2], Monique M.B. Breteler [1,3], and Folgerdiena M. de Vries [1]*

[1]Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1, Building 99, 53127 Bonn, Germany; [2]Institute of Clinical Pharmacology, Faculty of Medicine, RWTH, Aachen, Germany; and [3]Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Bonn, Germany

Received 3 April 2023; revised 21 July 2023; accepted 19 August 2023; online publish-ahead-of-print 25 August 2023

Despite increased emphasis on blood pressure control, hypertension is still widespread, and the quality of treatment is suboptimal.[1,2] The proportion of people taking antihypertensive medication whose blood pressure levels are higher than recommended is high.[1,2] However, less is known about the proportion of antihypertensive users whose blood pressure is lower than recommended. Because the harms of lowering systolic blood pressure (SBP) < 120 mmHg may outweigh the potential benefits, European guidelines currently advise against it.[3] A better understanding of the prevalence and determinants of over- and undertreatment of hypertension can help to improve the quality of care and identify individuals at risk of suboptimal treatment. The aim of this study was therefore to assess the prevalence and determinants of overtreatment and undertreatment of hypertension among antihypertensive users.

This cross-sectional study was based on the first 5000 participants (≥30 years) of the Rhineland Study, a population-based cohort study in Bonn, Germany. Sole exclusion criterion is insufficient command of the German language to provide written informed consent.

Blood pressure was measured with an oscillometric meter (three measurements at 10-min intervals; ≥5 min resting period; the mean of the last two measurements was used to calculate office blood pressure). Hypertension was defined as SBP of ≥140 mmHg and/or diastolic blood pressure (DBP) ≥ 90 mmHg,[3] and/or regular antihypertensive use. Medication data were collected by interview and included name, dosage, and current prescription status.[4] Antihypertensive users were categorized as *controlled* (SBP ≥ 120–<140 mmHg and DBP <90 mmHg), *overtreated* (SBP < 120 mmHg), or *undertreated* (SBP ≥ 140 and/or DBP ≥ 90 mmHg).

Multivariate logistic regression models were used to assess putative determinants (see Supplementary material online, *Table SA.1*) of over- and undertreated hypertension (reference: controlled users). All analyses were stratified by sex. Missing covariate data were imputed by predictive mean matching (10 bootstrap replicates). Statistical analyses were performed in R (version 4.1.1).

We excluded 196 participants because of incomplete blood pressure ($n = 129$) and medication ($n = 67$) data. Of the 4804 individuals included (mean age 55.0 years, range 30–95; 56.2% women), 1785 (37.2%; men: 42%, women: 33%, $P < 0.001$) had hypertension, of whom 1262 (70.7%) were taking antihypertensive medication (see Supplementary material online, *Table SA.2*). Among users, 46.7% ($n = 589$) had well-controlled blood pressure, 19.9% ($n = 251$) were overtreated, and 33.4% ($n = 422$) were undertreated (*Table 1*). There was no significant difference in the total defined daily dose (DDD) between the groups. Women were more likely to be overtreated than men (odds ratio [OR] 0.55, 95% confidence interval [CI] 0.39–0.76) (*Table 2*). The probability of overtreatment declined with increasing age in women (OR 0.95 per year, 95% CI 0.93–0.99), and men with chronic kidney disease (CKD) were more likely to be overtreated than men without CKD (OR 2.23, 95% CI 1.18–4.20). Furthermore, the odds of overtreatment decreased with increasing body mass index (BMI) (women: OR 0.96 per kg/m^2, 95% CI 0.92–1.00; men: OR 0.94 per kg/m^2, 95% CI 0.88–1.01). No sex difference was found among undertreated users. In women, the odds of undertreatment increased with age (OR 1.05 per year, 95% CI 1.02–1.08), and women with CKD were less likely undertreated than women without CKD (OR 0.50, 95% CI 0.29–0.88).

We observed a high prevalence of hypertension, and, similar to other studies, the prevalence was higher in men than in women.[2] However, inadequate treatment was more common in women (58% women; 49% men). This difference can be partly due to biological differences, such as the renin–angiotensin system and arterial stiffness.[5] Sex differences in the pharmacokinetics and pharmacogenetics of antihypertensives have been reported that cannot be attributed to differences in weight and body composition, as the sex dimorphism was still present after adjustment for those factors.[6] Nevertheless, current European guidelines on hypertension do not differentiate between women and men.[3,5]

Whether CKD patients are overtreated or intentionally set to low SBP cannot be conclusively answered. Nevertheless, the relationship between hypertension and CKD is complex and bidirectional and remains controversial.[7] Blood pressure usually increases with decreasing kidney function, and, in turn, a sustained increase in blood pressure accelerates renal disease, potentially motivating clinicians to aim for lower SBP, leading to overtreatment. Furthermore, younger people are more likely to be overtreated and older individuals more likely to be

* Corresponding author. Tel: +49 228 43302 943, Email: Dianna.deVries@dzne.de

Table 1 Characteristics of the study population

			Antihypertensive users ($n = 1262$)				
	All	Missing	Controlled	Overtreated	Undertreated	P^a	P^b
Participants, *n* (%)	4804 (100)		589 (46.7)	251 (19.9)	422 (33.4)	**<0.001**	**<0.001**
Age (years), *M* (SD)	55.0 (13.9)	0.0%	64.5 (11.8)	62.7 (11.6)	68.3 (10.8)	**0.042**	**<0.001**
Sex (women), *n* (%)	2702 (56.2)		284 (48.2)	152 (60.6)	233 (55.2)	**0.001**	**0.029**
Education, *n* (%)		0.8%					
Low	91 (1.9)		18 (3.1)	6 (2.4)	22 (5.3)	0.589	0.407
Middle	2141 (44.9)		325 (55.7)	129 (52.4)	241 (57.8)	Ref.	Ref.
High	2533 (53.2)		240 (41.2)	111 (45.1)	154 (36.9)	0.118	0.916
Smoking, *n* (%)		3.8%					
Never	2171 (47.0)		215 (37.9)	95 (40.1)	168 (41.3)	Ref.	Ref.
Former	1880 (40.7)		298 (52.5)	104 (43.9)	200 (49.1)	0.320	0.448
Current	572 (12.4)		55 (9.7)	38 (16.0)	39 (9.6)	0.112	0.636
BMI (kg/m^2), *M* (SD)	25.9 (4.5)	0.4%	28.4 (5.2)	27.6 (5.2)	27.5 (4.4)	**0.009**	0.081
SBP (mmHg), *M* (SD)	126.4 (16.1)	0.0%	129.7 (5.7)	112.0 (6.2)	152.5 (12.3)	**<0.001**	**<0.001**
DBP (mmHg), *M* (SD)	75.2 (9.4)	0.0%	74.7 (7.3)	67.7 (7.1)	83.3 (9.8)	**<0.001**	**<0.001**
Hypertension, *n* (%)	1785 (37.2)	0.0%	—	—	—	—	—
Duration of hypertension (years), *M* (SD)	12.9 (10.2)	0.5%	11.9 (10.1)	12.7 (10.3)	14.0 (11.7)	0.116	0.102
CVD, *n* (%)	504 (10.5)	0.3	154 (26.3)	70 (28.0)	110 (26.1)	0.098	0.190
CKD, *n* (%)	290 (6.3)	4.5%	89 (16.1)	52 (21.9)	72 (18.1)	**0.001**	0.239
Diabetes, *n* (%)	235 (5.0)	3.0%	84 (14.8)	32 (13.3)	51 (12.6)	0.943	0.272
Antihypertensive users, *n* (%)	1262 (26.3)	0.0%	—	—	—	—	—
Total DDD of antihypertensives[c], *M* (SD)	2.3 (1.9)	0.2%	2.3 (1.8)	2.3 (2.0)	2.4 (1.8)	0.865	0.894
Number of antihypertensives, *n* (%)		0.0%					
1	403 (31.9)		187 (31.7)	90 (35.9)	126 (29.9)	Ref.	Ref.
2	385 (30.5)		187 (31.7)	65 (25.9)	133 (31.5)	0.167	0.787
≥3	474 (37.6)		215 (36.6)	96 (38.2)	163 (38.6)	0.839	0.931
Users of lipid-lowering agents, *n* (%)	554 (11.5)	0.0%	176 (29.9)	78 (31.1)	138 (32.7)	0.090	0.967
Polypharmacy, *n* (%)	744 (15.5)	0.0%	244 (41.4)	118 (47.0)	200 (47.4)	**0.031**	0.822
HRT[d], *n* (%)	626 (13.1)	0.5%	47 (8.1)	34 (13.8)	36 (8.6)	0.473	0.756
Menopause[d], *n* (%)	1554 (57.5)	0.6%	246 (41.8)	118 (47.0)	215 (50.9)	0.769	0.342
Alcohol intake (g/day), *M* (SD)	20.0 (30.2)	11.5%	22.6 (29.2)	20.2 (31.0)	22.6 (34.1)	0.730	0.716
Global cognition (z-score), *M* (SD)	0.0 (0.6)	1.5%	−0.4 (0.6)	−0.3 (0.6)	−0.5 (0.6)	0.466	0.519

Data presented as mean ± standard deviation for quantitative variables/percentages for categorical variables. *P*-values in bold represent significant values. Group differences were assessed with logistic regression adjusted for age and sex (the variables age and sex were only adjusted for the other, respectively).
BMI, body mass index; CKD, chronic kidney disease; CVD, cardiovascular disease; DBP, diastolic blood pressure; DDD, defined daily dose; HRT, hormone replacement therapy (including oral contraceptives); *n*, number of participants; Ref., reference; SBP, systolic blood pressure.
[a]Adjusted for age and sex (overtreated compared to controlled).
[b]Adjusted for age and sex (undertreated compared to controlled).
[c]The corresponding calculation is explained in Supplementary material online, *Supplementary A.1*.
[d]Frequency among women only.

undertreated. Physicians are possibly more reluctant to target lower blood pressure levels in older than in younger patients, due to potential side effects or clear patient preferences.[8] Lastly, individuals with a low BMI are at increased risk of overtreatment. This is not surprising, as body weight is known to influence the pharmacokinetics of a drug.[9]

The adverse effects of undertreated or untreated hypertension are well known,[10] whereas the definition and consequences of overtreating hypertension is continuously debated. One study reported that lowering SBP < 130 mmHg had no beneficial effects,[11] while another study showed a lower relative risk of cardiovascular events.[12] Moreover, a J-shaped relationship between SBP/DBP and stroke incidence was reported in older people with hypertension, suggesting that 'the lower the better' may not be optimal, particularly in the elderly.[13] Additionally, overtreatment leads to high health care costs, and undertreatment burdens patients with treatment without clinical benefits.[14,15] Both cause avoidable harm and are important aspects of primary care that would benefit from systematic quality improvement.[15] Unsurprisingly, overtreatment and undertreatment have been described as 'the conjoined twins of modern medicine'.[16]

Strengths of our study include a broad age range, in-depth phenotyping, and high-quality data collection. We used interview-based medication data,[4] which may better reflect actual medication use than prescription data. Potential limitations, which are common in epidemiological studies, include our use of single blood pressure measurements and a possible healthy volunteer effect. However, the prevalence of hypertension and polypharmacy as well as the age and sex distribution

Table 2 Determinants of over- and undertreated hypertension (reference group: controlled antihypertensive users)

		All ($n = 1262$)			Women ($n = 669$)			Men ($n = 589$)		
Status	**Determinants**	**OR**	**95% CI**	***P***	**OR**	**95% CI**	***P***	**OR**	**95% CI**	***P***
Overtreated	Age[a] (per 1-year increase)	**0.98**	**0.96–0.99**	**0.**012	**0.95**	**0.93–0.99**	**0.**004	1.01	0.98–1.04	0.417
	Sex (men vs. women)	**0.55**	**0.39–0.76**	<0.001	—	—	—	—	—	—
	Education (low vs. middle)	0.77	0.29–2.02	0.589	1.02	0.37–2.83	0.972	—	—	—
	Education (high vs. middle)	1.31	0.94–1.83	0.108	1.27	0.79–2.04	0.319	1.20	0.72–1.98	0.484
	Smoking (former vs. never)	0.83	0.59–1.16	0.269	0.77	0.49–1.22	0.267	0.81	0.48–1.36	0.414
	Smoking (current vs. never)	1.33	0.81–2.18	0.255	1.27	0.65–2.46	0.484	1.69	0.74–3.85	0.209
	BMI (kg/m^2, per 1-unit increase)	**0.95**	**0.92–0.99**	**0.**004	**0.96**	**0.92–1.00**	**0.**037	0.94	0.88–1.01	0.077
	CVD (yes vs.no)	1.15	0.78–1.71	0.478	0.93	0.53–1.63	0.792	1.22	0.68–2.19	0.505
	CKD (yes vs.no)	**2.13**	**1.39–3.29**	**0.**001	1.50	0.80–2.80	0.203	**2.23**	**1.18–4.20**	**0.**014
	Diabetes (yes vs.no)	1.09	0.69–1.73	0.698	1.29	0.66–2.52	0.453	0.80	0.41–1.57	0.510
	Use of lipid-lowering agents (yes vs. no)	1.17	0.79–1.73	0.432	1.06	0.60–1.86	0.841	1.39	0.78–2.48	0.268
	Polypharmacy (yes vs. no)	1.29	0.89–1.88	0.183	1.39	0.84–2.29	0.197	1.32	0.71–2.46	0.379
	Used antihypertensives (2 vs. 1)	0.78	0.52–1.15	0.204	1.01	0.59–1.71	0.979	0.48	0.26–0.90	**0.**058
	Used antihypertensives (≥3 vs. 1)	0.97	0.65–1.43	0.866	1.46	0.86–2.47	0.157	0.63	0.34–1.16	0.135
	Alcohol consumption (g/day, per 1-unit increase)	1.00	0.99–1.00	0.574	1.00	0.99–1.01	0.690	1.00	0.99–1.01	0.731
	Global cognition (z-score, per 1-SD increase)	1.23	0.88–1.70	0.225	1.04	0.65–1.65	0.880	1.55	0.93–2.56	0.089
	Menopause[b] (yes vs. no)	—	—	—	0.80	0.37–1.75	0.581	—	—	—
	HRT[b] (yes vs. no)	—	—	—	1.11	0.64–1.92	0.713	—	—	—
Undertreated	Age (per 1-year increase)	**1.03**	**1.01–1.05**	**<0.001**	**1.05**	**1.02–1.08**	**0.001**	1.02	1.00–1.04	0.082
	Sex (men vs. women)	0.82	0.62–1.08	0.161	—	—	—	—	—	—
	Education (low vs. middle)	1.31	0.67–2.56	0.422	1.34	0.64–2.79	0.430	0.82	0.12–5.44	0.834
	Education (high vs. middle)	0.97	0.73–1.29	0.814	1.10	0.71–1.69	0.674	0.87	0.59–1.29	0.496
	Smoking (former vs. never)	0.87	0.66–1.15	0.341	0.81	0.55–1.20	0.293	1.03	0.68–1.56	0.880
	Smoking (current vs. never)	1.02	0.64–1.63	0.919	1.04	0.54–2.00	0.905	1.26	0.61–2.62	0.529
	BMI (kg/m^2, per 1-unit increase)	0.98	0.95–1.01	0.119	0.98	0.94–1.01	0.186	0.99	0.94–1.04	0.656
	CVD (yes vs. no)	0.78	0.56–1.08	0.136	0.81	0.51–1.31	0.392	0.73	0.46–1.17	0.189
	CKD[a] (yes vs. no)	0.88	0.61–1.27	0.489	**0.50**	**0.29–0.88**	**0.**016	1.18	0.70–1.99	0.533
	Diabetes (yes vs. no)	0.87	0.59–1.29	0.495	0.90	0.49–1.66	0.734	0.80	0.47–1.36	0.403
	Use of lipid-lowering agents (yes vs. no)	1.05	0.76–1.45	0.754	1.03	0.64–1.63	0.915	1.10	0.69–1.74	0.690
	Polypharmacy (yes vs. no)	1.13	0.83–1.54	0.448	1.29	0.84–1.99	0.246	0.96	0.58–1.56	0.857
	Used antihypertensives (2 vs. 1)	1.06	0.76–1.47	0.735	1.04	0.66–1.63	0.863	1.10	0.67–1.82	0.698
	Used antihypertensives (≥3 vs. 1)	1.09	0.78–1.53	0.610	0.96	0.61–1.53	0.874	1.26	0.76–2.10	0.375
	Alcohol consumption (g/day, per 1-unit increase)	1.00	1.00–1.00	0.712	1.00	0.99–1.01	0.844	1.00	0.99–1.00	0.804
	Global cognition (z-score, per 1-SD increase)	1.01	0.77–1.32	0.954	1.14	0.76–1.70	0.534	0.86	0.59–1.27	0.455
	Menopause[b] (yes vs. no)	—	—	—	0.59	0.25–1.37	0.217	—	—	—
	HRT[b] (yes vs. no)	—	—	—	0.92	0.55–1.53	0.740	—	—	—

Values in bold represent significant values ($P < 0.05$).
BMI, body mass index; CI, confidence intervals; CKD, chronic kidney disease; CVD, cardiovascular disease; HRT, hormone replacement therapy (including oral contraceptives); *n*, number of participants; OR, odds ratio.
[a]Significant sex interaction term ($P < 0.05$).
[b]In women only.

in our study are comparable to the German population.[17] Our population might be more health conscious, which could affect the quality of treatment, but this would mean that the quality of treatment in the general population may be even worse. Finally, the generalizability of our findings to other countries might be limited by differences in health care systems.

In conclusion, we found that antihypertensive treatment was suboptimal in more than half of the individuals, and women were more likely to receive suboptimal treatment than men. A one-size-fits-all approach seems inadequate, and hypertension guidelines should emphasize the avoidance of overtreatment through deintensification.

Authors' contribution

N.A.: conceptualization, methodology, formal analysis, writing—original draft, writing—review and editing, and visualization. G.P.: conceptualization, data curation, and writing—review and editing. J.C.S.: conceptualization, methodology, writing—review and editing, and supervision. M.M.B.B.: conceptualization, methodology, resources, data curation, writing—review and editing, supervision, and funding acquisition. F.M.d.V.: conceptualization, methodology, data curation, writing—review and editing, visualization, supervision, and project administration.

Ethics approval statement and informed consent

Approval to undertake the study was obtained from the ethics committee of the University of Bonn, Medical Faculty. The study was carried out in accordance with the recommendations of the International Council for Harmonisation (ICH) Good Clinical Practice (GCP) standards (ICH-GCP). We obtained written informed consent from all participants in accordance with the Declaration of Helsinki. Protocol number: main study: Lfd.Nr. 338/15.

Supplementary material

Supplementary material is available at *European Journal of Preventive Cardiology*.

Funding

This work was supported by a grant provided by the Federal Institute for Drugs and Medical Devices (Bundesinstitut für Arzneimittel und Medizinprodukte), Bonn, Germany (V-15646/68502/2014-2016 and V-17746/68502/2018-2020). The funding sources have not been involved in study design; in the collection, analysis, and interpretation of the data; in the writing; or in the decision to submit the article for publication.

Conflict of interest: None declared.

Data availability

The data sets for this manuscript are not publicly available because of data protection regulations. Access to data can, however, be provided to scientists in accordance with the Rhineland Study's Data Use and Access Policy. Requests to access the data sets should be directed to M.M.B.B., RS-DUAC@dzne.de.

References

1. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet* 2021; **398**:957–980.
2. Beaney T, Schutte AE, Tomaszewski M, Ariti C, Burrell LM, Castillo RR, *et al.* May Measurement Month 2017: an analysis of blood pressure screening results worldwide. *Lancet Glob Heal* 2018;**6**:e736–e743.
3. G. M, G. DB, A. D et al. 2018 Guidelines for the management of arterial hypertension: the Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). 2018.
4. Alaeddin N, Stingl JC, Breteler MMB, Vries FM. Validation of self-reported medication use applying untargeted mass spectrometry-based metabolomics techniques in the Rhineland study. *Br J Clin Pharmacol* 2022;**88**:2380–2395.
5. Cadeddu C, Franconi F, Cassisa L, Campesi I, Pepe A, Cugusi L, *et al.* Arterial hypertension in the female world. *J Cardiovasc Med* 2016;**17**:229–236.
6. Kalibala J, Pechère-Bertschi A, Desmeules J. Gender differences in cardiovascular pharmacotherapy—the example of hypertension: a mini review. *Front Pharmacol* 2020;**11**:564.
7. Hamrahian SM, Falkner B. Hypertension in chronic kidney disease. In: Islam MS (ed.), *Hypertension: From Basic Research to Clinical Practice*. Cham: Springer International Publishing; 2017. p307–325.
8. van Middelaar T, Ivens SD, van Peet PG, Poortvliet RKE, Richard E, Pols AJ, *et al.* Prescribing and deprescribing antihypertensive medication in older people by Dutch general practitioners: a qualitative study. *BMJ Open* 2018;**8**:e020871.
9. Cheymol G. Effects of obesity on pharmacokinetics implications for drug therapy. *Clin Pharmacokinet* 2000;**39**:215–231.
10. Zhou D, Xi B, Zhao M, Wang L, Veeranki SP. Uncontrolled hypertension increases risk of all-cause and cardiovascular disease mortality in US adults: the NHANES III Linked Mortality Study. *Sci Rep* 2018;**8**:9418.
11. Mancia G, Kjeldsen SE, Zappe DH, Holzhauer B, Hua TA, Zanchetti A, *et al.* Cardiovascular outcomes at different on-treatment blood pressures in the hypertensive patients of the VALUE trial. *Eur Heart J* 2016;**37**:955–964.
12. Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, *et al.* Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet* 2016;**387**:957–967.
13. Vokó Z, Bots ML, Hofman A, Koudstaal PJ, Witteman JC, Breteler MM. J-shaped relation between blood pressure and stroke in treated hypertensives. *Hypertension* 1999;**34**:1181–1185.
14. Lyu H, Xu T, Brotman D, Mayer-Blackwell B, Cooper M, Daniel M, *et al.* Overtreatment in the United States. *PLoS One* 2017;**12**:e0181970.
15. Kearney M, Treadwell J, Marshall M. Overtreatment and undertreatment: time to challenge our thinking. *Br J Gen Pract* 2017;**67**:442–443.
16. Heath I. Role of fear in overdiagnosis and overtreatment—an essay by Iona Heath. *BMJ* 2014;**347**:g6123–g6123.
17. de Vries FM, Stingl JC, Breteler MMB. Polypharmacy, potentially inappropriate medication and pharmacogenomics drug exposure in the Rhineland Study. *Br J Clin Pharmacol* 2021;**87**:2732–2756.

SUPPLEMENTARY

A.1 Calculation of total Defined Daily Dose (DDD) of antihypertensives

To investigate the dose, we used the DDD, which is taken to be the average maintenance dose per day for a drug used for its main indication in adults. As the DDD provides a fixed unit of measurement, it allows comparability and dose homogenisation. For each antihypertensive drug and for each participant, we divided the consumed daily dose (mg/day) by the DDD (mg/day), provided by the World Health Organization (WHO), and then added the quotients to determine the total DDD for each participant.

Table A.1 Definition of demographic and clinical characteristics

	Characteristic	Definition
General characteristics	Age	Age range 30-95 years, mean-centered
	Sex	Women, men
	Education	Based on the International Standard Classification of Education 2011 (ISCED): low (lower secondary education or below), middle (upper secondary education to undergraduate university level), high (postgraduate university study)
	Smoking	Never, former and current smokers
	Body mass index (BMI)	Body mass divided by the square of the body height (kg/m^2)
Comorbidities	Cardiovascular disease (CVD)	Based on the self-reported presence of one or more of these conditions: myocardial infarction, coronary artery disease, cardiac insufficiency, cardiac pacemaker, peripheral artery disease, stroke, surgery on large vessels such as aortic, carotid or peripheral vessel, heart valve disease
	Chronic kidney disease (CKD)	Estimated glomerular filtration rates based on cystatin c levels (no CKD ≥60 ml/min/1.73m^2; CKD <60ml/min/1.73m^2
	Diabetes	Self-reported physician diagnosis, and/ or glycated haemoglobin (HbA_{1c}) (no diabetes <6.5%; diabetes ≥6.5%), fasting glucose (no diabetes <126 mg/dl; diabetes ≥126 mg/dl) measured in fasting morning blood, and/ or antidiabetic medication use (ATC: A10)
Medication	Number of used antihypertensive drugs	Use of 1, 2 or ≥3 antihypertensive drugs (ATC: C02, C03, C07, C08, C09)
	Polypharmacy	Regular use of ≥5 prescribed drugs
	Lipid-lowering agents	Use of lipid-lowering agents (ATC: C10)
Nutrition	Alcohol intake	Dietary intake of alcohol (g/day) based on information from food intake frequency questionnaires
Cognition	Global cognition (z-score)	Derived from a cognitive test battery assessing episodic verbal memory, working memory, executive function and processing speed
Women-specific characteristics	Menopause	Self-reported information on premenopausal and postmenopausal status

Table A.2 Number of used antihypertensives (a), most frequently used antihypertensive drug classes (b) and most frequent dual (c)/ triple combinations (d)

a. Number of used antihypertensive drugs				**b. Most frequently used antihypertensive drug classes**			
	n	**%**	**95% CI**		**n**	**%**	**95% CI**
1 drug	403	31.9	(29.4; 34.6)	ARBs	581	46.0	(43.3; 48.8)
2 drugs	385	30.5	(28.0; 33.1)	β-blockers	501	39.7	(37.0; 42.5)
3 drugs	290	23.0	(20.7; 25.4)	Diuretics	451	35.7	(33.1; 38.4)
4 drugs	142	11.3	(9.6; 13.1)	ACE-inhibitors	380	30.1	(27.6; 32.7)
5 drugs	42	3.3	(2.4; 4.5)	CCBs	323	25.6	(23.2; 28.1)
c. Most frequent dual therapy combinations				**d. Most frequent triple therapy combinations**			
	n	**%**	**95% CI**		**n**	**%**	**95% CI**
ARBs + Diuretics	118	30.6	(26.1; 35.5)	Diuretics + ARBs + CCBs	60	20.7	(16.2; 25.8)
ACE- inhibitors + Diuretics	66	17.1	(13.5; 21.3)	β-blockers + Diuretics + ARBs	53	19.3	(14.9; 24.3)
ARBs + CCBs	53	13.8	(10.4; 17.6)	β-blockers + Diuretics + ACE-inhibitors	32	11.0	(7.7; 15.2)
β-blockers + ARBs	49	12.7	(9.6; 16.5)	ACE-inhibitors + Diuretics + CCBs	15	5.2	(2.9; 8.4)
β-blockers + ACE-inhibitors	46	11.9	(8.9; 15.6)	β-blockers + ACE-inhibitors + CCBs	13	4.5	(2.4; 7.5)

Abbreviations: ACE-Inhibitors, Angiotensin-converting-enzyme inhibitors; ARBs, Angiotensin II receptor blockers; CCBs, calcium channel blockers; CI, confidence interval; n, number of participants

3.4 The impact of proton pump inhibitors on cognition and brain structure in the general population: the Rhineland Study (*submitted manuscript*)

This is a non-final version submitted to the journal *Neurology* (December 2023)

The impact of proton pump inhibitors on cognition and brain structure in the general population: the Rhineland Study

Nersi Alaeddin [a], Alexandra Koch [a], Rika Etteldorf [a], Julia C. Stingl [b], Monique M.B. Breteler [a, c], Folgerdiena M. de Vries [a]

a. *Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany*
b. *Institute of Clinical Pharmacology, Faculty of Medicine, RWTH Aachen, Germany*
c. *Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Bonn, Germany*

Corresponding author:
Folgerdiena M. de Vries, PhD
Deutsches Zentrum für Neurodegenerative Erkrankungen e.V. (DZNE)
Venusberg-Campus 1, Building 99
53127 Bonn, Germany
Email: Dianna.deVries@dzne.de
Tel: +49-228-43302-943

Abstract

Background and Objectives: The current literature on proton pump inhibitors (PPIs) and their impact on cognition and dementia presents conflicting results and lacks exploration of early indicators of cognitive decline through microstructural brain changes. This study aims to address these gaps by exploring associations between PPI use and cognitive function, volumetric brain measures, and microstructural diffusion tensor imaging (DTI) measures.

Methods: We used cross-sectional data from the Rhineland Study, a population-based cohort study in Bonn, Germany. Medication data were collected by interview and magnetic resonance imaging (MRI) was conducted using a 3T MRI scanner to assess structural volumes with the Freesurfer processing pipeline. White matter microstructure was assessed through diffusion-weighted MRI. Using multivariate linear regression, we investigated the relationship of PPI use with cognition and brain macro- and microstructural measures (fractional anisotropy (FA) and mean diffusivity (MD) as measures of white matter integrity). We stratified the analyses by duration of PPI use (short-term use: <3 years, long-term use: ≥3 years), and the analyses of cognitive outcomes were additionally stratified by age (<65 years; ≥65 years).

Results: The study included 7,480 eligible participants (mean age 55.4±13.7 years, range 30-95 years, 56.6% women). Compared to non-users, PPI users exhibited poorer global cognition (β= -0.08, 95%CI -0.14; -0.01, *p*=.020), total memory (β=-0.09, 95%CI -0.17; -0.01, *p*=.029), and working memory (β= -0.13, 95%CI -0.22; -0.03, *p*=.009) among younger individuals. Specifically, in the subset of younger long-term PPI users, PPI use was associated with worse total memory (β= -0.12, 95%CI -0.23; -0.01, *p*=.030) and working memory (β= -0.14, 95%CI -0.26; -0.01, *p*=.032). Notably, no significant associations were found between PPI use and volumetric brain measures or fractional anisotropy. Instead, PPI users exhibited higher MD in specific brain regions associated with cognitive function.

Discussion: Our findings indicate that prolonged PPI use, especially in younger individuals, is associated with poorer cognitive performance and potential white matter integrity disruptions. Further research is necessary to explore whether this relationship is causal and which mechanisms underlie the relation between PPI use and cognitive decline. Our findings emphasize the need for caution in PPI prescription due to their potential adverse effects.

1. INTRODUCTION

Proton pump inhibitors (PPIs) are commonly prescribed drugs for acid-related gastric disorders that are widely used in Europe and the USA [1,2]. PPIs are available both by prescription and over-the-counter [3,4]. A German study using health insurance data revealed that prescription rates almost doubled between 2005 and 2013 (from 8.2% to 16.2%). This trend showed distinct sex and age disparities, with women using PPIs more than men and older people more than younger people [5]. Evidently, in Germany, PPIs have claimed the 11th place among the most prescribed drugs, with over 5 million prescriptions in 2019 [6]. Approximately 40% of these prescriptions are estimated to be inappropriate due to use of PPIs longer than recommended, elevated risk-benefit ratios, and drug-drug interactions [7–9]. This is particularly worrisome, as mounting evidence challenges the presumed safety of long-term PPI use [10], linking them to various long-term adverse effects, including increased susceptibility to bacterial infections, pneumonia, cardiovascular disease, vitamin deficiencies, bone fractures, chronic kidney disease and even dementia [11–13].

While there is limited evidence linking PPIs to cognition and dementia via biological mechanisms. several hypotheses, although not tested in humans, suggest potential associations with increased risk of dementia: (I) In murine models, PPIs shift the cleavage site of amyloid-β precursor protein (APP), resulting in more β-amyloid$_{42}$ and less β-amyloid$_{38}$ [35]; (II) PPIs contribute to a less acidic microglial environment, potentially posing a risk for Alzheimer's disease by reducing β-amyloid clearance [36]; (III) PPIs inhibit choline acetyltransferase, thereby affecting acetylcholine biosynthesis [37]. Finally, PPIs reduce the gastric acid needed for vitamin B12 absorption [38], which has been suggested to contribute to neurodegeneration and cognitive impairment [39].

Numerous studies have investigated associations between PPIs and dementia, with mixed findings. Some studies suggest an increased risk of dementia with PPI use [14–20], while others found no discernible risk [21–25] or, intriguingly, even a reduced risk [26–28]. These conflicting results could be attributed to methodological limitations, including protopathic bias, where the drug is prescribed because of early symptoms of an undiagnosed condition, potentially resulting in erroneous causal inferences between exposure and outcome. Given the long prodromal phase of dementia, this bias holds particular relevance for studies linking PPI use to dementia risk [29,30]. Moreover, the association between PPI use and cognitive decline also remains controversial. While a clinical trial involving sixty young volunteers reported negative effects of PPI use on cognition, other prospective and population-based studies found no such associations [31–34]. In the population-based SHIP study, PPI users scored lower on memory tests, but no association was found between PPI use and brain volume or age [34]. However, this study did not

distinguish between short-term and long-term users, potentially biasing effect estimates. Additionally, the study did not examine the effects of PPIs on microstructural brain parameters.

However, microstructural brain imaging through diffusion tensor imaging (DTI) shows promise in detecting early structural changes indicative of cognitive decline. DTI assesses the diffusion rates of water through tissue, revealing hidden variations in tissue integrity that are not visible on conventional MRI scans [40]. The use of imaging techniques that can depict these very early, preclinical changes holds great promise to assess associations between PPI use and brain function. Hence, this study aims to investigate associations between PPI use and cognition, volumetric brain measures, and microstructural DTI measures in the general population.

2. METHODS

2.1 Study design & setting

The Rhineland Study is an ongoing prospective community-based cohort study that started recruitment in 2016. All residents aged ≥30 years from two geographically defined areas in Bonn, Germany, are invited to participate in this single-center study. The municipality provides contact details of eligible participants. Participation is by invitation only and independent of health status. The only exclusion criterion is insufficient knowledge of the German language to give written informed consent in accordance with the Declaration of Helsinki. Participants will be followed for decades, with follow-up examinations every three to four years. All participants undergo in-depth phenotyping, including assessment of cardiovascular health, cognitive testing, MRI scans, neurological function and medication use. The study was approved by the Ethics Committee of the University of Bonn, Medical Faculty.

2.2 Study population

As recruitment to the Rhineland Study is ongoing, we used data from all participants who completed baseline examinations between March 2016 and November 2021 (n=8,318) for this cross-sectional analysis. Of these, we excluded 838 participants because of missing data on PPI exposure (n=125), as needed PPI use (n=533), stroke (n=122), reported Parkinson's disease (n=57), or traumatic brain injury (n=1). For the analyses regarding cognition outcomes, we additionally excluded participants without cognitive data (n=297) leaving 7,183 participants, and for the analyses using MRI outcome measures we additionally excluded participants without MRI data (n=2,506) leaving 4,974 participants (**Figure 1**).

2.3 Medication data collection and PPI use

Participants were asked to bring the original packages of all medications (including over-the-counter drugs and excluding homeopathic drugs) and prescribed dietary supplements that they were currently taking

and had taken as needed in the past year. We registered the name, dosage and Anatomical Therapeutic Chemical (ATC) code of each medication and supplement.

PPI use was defined as regular use of medications starting with ATC code A02BC. Participants who used PPIs on an as-needed basis were excluded from the analyses. Participants who reported not using PPIs were defined as non-users. In addition, PPI users were divided into short-term users (<3 years) and long-term users (≥3 years).

2.4 Cognitive assessment

A battery of cognitive tests was administered according to standardized protocols in German. Details of the cognitive battery are given in **Appendix A.**

Scores were calculated for different cognitive domains: processing speed, executive function, working memory and episodic verbal memory. Processing speed was assessed using the Trail-making Test A [41] (time to task completion) and the eye-tracking Pro-saccade task (mean saccadic latency, i.e. time needed to initiate a saccade). Executive function was assessed using a categorical Word Fluency Task (total number of animals named), the Trail-making Test B (time to task completion) and the eye-tracker Anti-saccade task (percentage of direction errors). Working memory was assessed with a Corsi block-tapping Test [42] and a digit span test (maximum forward and backward span). Episodic verbal memory was assessed using the total immediate recall (sum over trials 1-5) and the delayed recall on the Verbal Learning and Memory Test (VLMT) [43,44]. For the cognitive domains, the individual standardized scores within the domains were averaged to calculate the composite score.

An overall memory score was then calculated by averaging the working memory and episodic verbal memory scores, and a global cognitive score was calculated from the average of all cognitive domain scores.

2.5 Magnetic resonance imaging

Magnetic resonance imaging (MRI) was performed on eligible participants using a dedicated 3 Tesla MRI scanner (MAGNETOM Prisma, Siemens Healthineers, Erlangen, Germany) equipped with an 80 mT/m gradient system and a 64-channel phased-array head-neck coil. We assessed the effects of PPI use on total brain volume, cortical grey matter volume, white matter (WM) volume, ventricle volume, hippocampal volume, and cortical thickness. Structural volumes and thicknesses were determined using the standard Freesurfer processing pipeline (http://surfer.nmr.mgh.harvard.edu/) on T1-weighted MR images [45,46].

We further assessed the effect of PPI use on WM microstructure by using diffusion-weighted MRI (dMRI). DMRI probes brain tissue microstructure by measuring water diffusion properties. We examined fractional anisotropy (FA) and mean diffusivity (MD) obtained by fitting the diffusion tensor model to dMRI scans using the MDT framework. FA is a scalar metric ranging from 0 to 1, with higher values indicating higher

anisotropy, i.e. directionality, of water diffusion which is restricted within the complex tissue microstructure in the brain WM. MD is a scalar measure in units of mm^2/s and reflects the extend of water diffusion, with higher values indicating higher water mobility, i.e. less restriction of water movement within the WM microstructure. Age-related changes in WM microstructure are reflected by decreased FA and increased MD, typically summarized as a decline in WM integrity [47–49]. We examined the average FA and MD values for the entire brain and specifically in the following regions of interest within the brain's white matter, which have been associated to cognition [50,51]: Body of corpus callosum, left/right cingulum (cingulate gyrus), left/right cingulum (hippocampus), left/right corticospinal tract, fornix (column and body), genu of corpus callosum, left/right posterior thalamic radiation, splenium of corpus callosum, left/right superior longitudinal fasciculus, left/right sagittal stratum (including inferior longitudinal and fronto-occipital fasciculus) and left/right uncinate fasciculus. Further details on MRI acquisition and processing are provided in **Appendix B.**

2.6 Confounders

The association between PPI use and cognitive and brain measures may be confounded by participant characteristics. Therefore, we simultaneously adjusted for the following characteristics, based on biological plausibility: (I) demographics, including: age (mean-centered), sex, education (based on International Standard Classification of Education (ISCED-11); categorization: low, middle, high), self-reported first-language (native German, non-native), smoking status (never, former, current), body mass index (BMI); (II) comorbidities, including: hypertension (based on current regular use of antihypertensives, and/or mean systolic blood pressure ≥140 and/or mean diastolic blood pressure ≥90) and diabetes (based on current use of antidiabetic drugs, and/ or Hba1c (no diabetes <6.5%; diabetes ≥6.5%) and/ or fasting glucose (no diabetes <126 mg/dL; diabetes ≥126 mg/dL) measured in morning fasting blood); (III) medication use, including: anticholinergic medication (based on Anticholinergic Burden (ACB) score) [52], antidepressants, antithrombotic medications, statins, and nonsteroidal anti-inflammatory drugs (NSAIDs); (IV) perceived stress through the Perceived Stress Scale (PSS) [53].

2.7 Statistical analysis

Descriptive statistics were used to summarize participant characteristics. Group differences were calculated using chi-squared (categorical variables) and ANOVA tests (continuous variables), and separately adjusted for age using logistic regression.

We examined the relation between PPI use and cognitive and brain outcomes using multivariable linear regression. All outcome measures were z-standardized. In the analysis of cognitive domains, we simultaneously adjusted for age, age^2 (mean-centered to avoid collinearity), sex, education, and the confounders described in section 2.6 as independent variables. In addition, executive function and

episodic verbal memory were adjusted for participants' self-reported first language, as the estimated scores of these two domains are strongly influenced by language proficiency. As age and PPI use yielded a significant interaction term in all cognitive domains, we conducted separate analyses for younger (<65 years) and older (≥65 years) individuals. Analyses were then performed comparing long-term (≥3 years) and short-term (<3 years) PPI use with non-use.

The same approach was used for the brain outcomes, with additional correction for total intracranial volume for volumetric brain outcomes. Here, no age stratification was performed, as there was no significant age interaction term in any of the brain measures. Because we were interested in the individual relations between PPI use and cognitive and brain outcomes, rather than testing a joint hypothesis, we did not correct for multiple testing [54]. All analyses were conducted in RStudio version 4.1.1.

3. RESULTS

3.1 Study population

The characteristics of PPI users and non-users are shown in **Table 1**. Included participants (n=7,480; mean age: 55.4 ±13.7 years, range 30-95 years; 56.6% women) were younger (p<.001) than excluded participants (n=838; mean age: 60.7 ±14.2 years, range 30-91 years; 54.7% women). Overall, 5.7% (n=426) of the study population were regular PPI users, with more women than men (53.4%) using PPIs. The majority reported long-term (≥3 years) PPI use (60.6%) and PPI users were older than non-users (65.3 vs. 54.7 years; p<.001). Pantoprazole (n=319) was the most commonly used PPI, followed by omeprazole (n=73), esomeprazole (n=21), lansoprazole (n=9) and rabeprazole (n=4).

3.2 Cognition

Across all cognitive domains, we observed a consistent but non-significant pattern of poorer cognitive performance in users of PPIs compared to non-users. Stratified analyses showed that PPI use was significantly associated with worse global cognition (β= -0.08, 95%CI -0.14; -0.01, p=.020), and total memory (β=-0.09, 95%CI -0.17; -0.01, p=.029), and working memory performance (β= -0.13, 95%CI -0.22; -0.03, p=.009), in younger individuals (aged <65 years) (**Table 2**). The effect sizes were comparable to an average age-related decline in global cognition, total memory and working memory, of 1.5, 2 and 7.5 years, respectively.

In both younger (<65 years) and older (≥65 years) individuals, no significant differences in cognitive performance were observed between short-term PPI users and non-users. However, among younger long-term PPI users, we observed significant negative effects of PPI use, resulting in worse total memory (β= -0.12, 95%CI -0.23; -0.01, p=.030) and working memory performance (β= -0.14, 95%CI -0.26; -0.01, p=.032) compared to non-users. Although the effect size for working memory in young short-term users

is comparable to long-term users (β=-0.12 vs. β=-0.14), no significance was found (β= -0.12, 95%CI -0.25; -0.02, *p*=.101).

3.3 Macro- and microstructural brain measures

We identified no significant differences in any of the assessed brain macrostructural measures (**Table 3**), nor with global fractional anisotropy and the investigated region-specific FA parameters (data not shown) between PPI users and non-users.

We observed higher global MD (β= 0.11, 95%CI 0.0; 0.23, *p*=.048), elevated MD in the body of the corpus callosum (β= 0.20, 95%CI 0.08; 0.33, *p*=.002) and in the left cingulum (cingulate gyrus) (β= 0.13, 95%CI 0.00; 0.25, *p*=.048) in PPI users compared to non-users.

In subgroup analyses, short-term users showed higher MD in the body of the corpus callosum (β= 0.20, 95%CI 0.01; 0.39, *p*=.037) and the left cingulum (hippocampus) (β= 0.21, 95%CI 0.01; 0.41, *p*=.042) compared to non-users (**Table 4**).

In long-term users, MD was higher in the body of the corpus callosum (β= 0.20, 95%CI 0.04; 0.36, *p*=.014) and the right corticospinal tract (β= 0.19, 95%CI 0.01; 0.37, *p*=.036) compared to non-users.

4. DISCUSSION

In the context of the population-based Rhineland Study, we investigated the impact of PPIs on cognitive function, and macro- and microstructural brain measures. We observed that younger PPI users, particularly does with a longer duration of use, had poorer cognitive performance compared to non-users. While volumetric brain measures showed no discernible differences between PPI users and non-users, we observed a higher mean diffusivity among PPI users in some brain regions, especially the corpus callosum.

Younger PPI users exhibited poorer performance across various cognitive domains, namely global cognition, total memory and working memory, when compared to non-users. These findings align with a small clinical trial conducted by Akter et al. on 60 young participants, reporting impaired cognitive performance in visual memory, attention, executive function and working/ planning function [31]. Further, Ahn and colleagues reported lower VLMT scores in PPI users [34]. In contrast, we did not find a significant association between PPI use and episodic verbal memory. Instead, the most pronounced effect was identified in working memory, which likely also drives the observed association between PPI use and the total composite memory score.

Moreover, we observed poorer cognitive performance in long-term PPI users (<65 years). This observation aligns with stratified analyses conducted in two age groups of a Danish middle-aged cohort,

where a stronger negative effect was seen in participants <57 years, even though statistical significance was not reached, likely due to the limited number of PPI users in these subsets [33]. The absence of noticeable effects in older participants may be due to competing causes like comorbidities, which can have a more substantial impact on cognitive performance. Another contributing factor could be the increased sensitivity of younger individuals to medications, potentially linked to age-related changes in pharmacokinetics and pharmacodynamics of medications [55].While there is a potential age-related influence of PPIs on cognitive function, causality remains uncertain.

Several hypotheses link cognitive decline and dementia to biological mechanisms [35–37]. Yet, the relevance of these hypotheses to human health remains uncertain due to lack in human testing. A more accepted theory connects cognitive decline with vitamin B12 deficiency [39], supported by a case-control study showing an association between gastric acid inhibitors and reduced vitamin B12 absorption [38]. Given PPIs' impact on vitamin B12 absorption, it is plausible that they can potentially contribute to cognitive decline.

We did not find any significant associations between PPI use and brain volumetric measures, consistent with prior research [34]. However, we observed an association of PPI use with elevated mean diffusivity in certain brain regions linked to cognitive function. Short-term users had higher MD in specific brain regions, such as in the body of the corpus callosum, left cingulum (hippocampus), and left uncinate fasciculus, compared to non-users. Long-term users similarly displayed higher MD, particularly in the body of the corpus callosum and the right corticospinal tract. These findings align with our cognitive performance observations in PPI users, given that the aforementioned regions are known to be associated with cognitive function [50]. Additionally, previous research has indicated that increased MD reflects decreased WM microstructure, a phenomenon often observed in the ageing brain or in brains affected by diseases [47–49]. Although our findings also hint at a potential link between PPI use and cognitive impairment [56–58], it is important to note that our results do not conclusively identify WM as the sole critical factor, necessitating further investigations.

Changes in PPI exposure during the prodromal phase of dementia could falsely show a negative effect of PPI use on cognition and dementia risk. This phase often involves prescription of medications to manage symptoms like depression, anxiety, sleep disturbances and cognitive impairment [59]. Some of these prescribed medications themselves can induce dyspeptic symptoms, for which PPIs are commonly recommended. Additionally, depression and anxiety may themselves cause dyspeptic symptoms [60], which may lead to PPI use. To address and mitigate these complexities, we implemented several approaches. First, in addition to controlling for a variety of confounding factors, we conducted subgroup

analyses within two age groups (younger (<65 years) and older individuals (≥65 years) for the cognitive outcomes. This stratification allowed us to explore potential age-related variations in the effects of PPI use. Furthermore, we differentiated between short-term and long-term PPI use, uncovering duration-dependent impacts that added depth to our analyses.

The strength of our study also lies in its ability to assess the effects of PPI use on both macrostructural and, notably, microstructural measures of the brain, representing a novel contribution to the field. Given the multifactorial potential influences on cognitive and brain function, our detailed phenotyping of our participants enabled us to control for numerous confounding factors. Furthermore, we gathered medication data through interviews, capturing medication information, including over-the-counter purchased PPIs, which enhances the comprehensiveness of our study. The accuracy and reliability of medication data collected in the Rhineland Study have been published elsewhere [61]. However, we must acknowledge certain limitations in our study. Its cross-sectional nature restricted our ability to explore temporal associations, leaving us unable to confirm whether PPI use preceded the observed changes in cognitive performance and brain microstructural measures. Therefore, we cannot definitively rule out the possibility of reversed causality. Furthermore, as in all pharmacoepidemiological studies, the potential for residual confounding by indication, particularly among long-term users, necessitates consideration. Despite our extensive adjustments, the influence of unmeasured variables on our findings cannot be entirely ruled out.

To conclude, our study reveals negative effects of PPI use on cognition, especially in younger long-term users. We did not find significant associations between PPI use and brain volume, but observed higher mean diffusivity in PPI users in specific brain regions, indicating potential white matter integrity impacts. Highlighting the imperative for deprescribing PPIs, especially in younger individuals, future studies necessitate larger sample sizes and longitudinal data for definitive conclusions. Meanwhile, exercising caution to PPI use is advisable to avoid potential risks of cognitive decline and changes in brain structure.

5. Contributors

Nersi Alaeddin: Conceptualization, Methodology, Formal analysis, Writing- Original Draft, Writing- Review & Editing, Visualization;

Alexandra Koch: Data Curation, Writing- Review & Editing;

Rika Etteldorf: Data Curation, Writing- Review & Editing

Julia C. Stingl: Conceptualization, Methodology, Writing- Review & Editing, Supervision;

Monique M.B. Breteler: Conceptualization, Methodology, Resources, Data Curation, Writing- Review & Editing, Supervision, Funding Acquisition;

Folgerdiena M. de Vries: Conceptualization, Methodology, Data Curation, Writing- Review & Editing, Visualization, Supervision, Project Administration.

6. Study funding

This work was supported by a grant provided by the Federal Institute for Drugs and Medical Devices (BfArM), Bonn, Germany (V-15646/68502/2014-2016 and V-17746/68502/2018-2020).

The funding sources have not been involved in study design; in the collection, analysis and interpretation of the data; in the writing or in the decision to submit the article for publication.

7. Declaration of interest

None.

8. Data availability statement

The datasets for this manuscript are not publicly available because of data protection regulations. Access to data can, however, be provided to scientists in accordance with the Rhineland Study's Data Use and Access Policy. Requests to access the datasets should be directed to Prof. Dr. Dr. Monique M.B. Breteler, RS-DUAC@dzne.de.

9. References

1. Lanas A. We Are Using Too Many PPIs, and We Need to Stop: A European Perspective. *Am J Gastroenterol* 2016;**111**:1085–6.

2. Schumock GT, Li EC, Suda KJ *et al*. National trends in prescription drug expenditures and projections for 2016. *Am J Heal Pharm* 2016;**73**:1058–75.

3. Heidelbaugh JJ, Kim AH, Chang R *et al.* Overutilization of proton-pump inhibitors: what the clinician needs to know. *Therap Adv Gastroenterol* 2012;**5**:219–32.

4. Johnson DA, Katz PO, Armstrong D *et al.* The Safety of Appropriate Use of Over-the-Counter Proton Pump Inhibitors: An Evidence-Based Review and Delphi Consensus. *Drugs* 2017;**77**:547–61.

5. Hoffmann F, Glaeske G, Schmiemann G. Steigende Verordnungszahlen von Protonenpumpenhemmern im ambulant-ärztlichen Bereich über die Jahre 2005 − 2013. *Z Gastroenterol* 2014;**53**:95–100.

6. Schwabe U, Ludwig W-D. Arzneiverordnungen 2019 im Überblick. *Arzneiverordnungs-Report 2020*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2020, 3–41.

7. Pasina L, Nobili A, Tettamanti M *et al.* Prevalence and appropriateness of drug prescriptions for peptic ulcer and gastro-esophageal reflux disease in a cohort of hospitalized elderly. *Eur J Intern Med* 2011;**22**:205–10.

8. Heidelbaugh JJ, Goldberg KL, Inadomi JM. Magnitude and economic effect of overuse of antisecretory therapy in the ambulatory care setting. *Am J Manag Care* 2010;**16**:e228-34.

9. Yang Y, Metz DC. Safety of Proton Pump Inhibitor Exposure. *Gastroenterology* 2010;**139**:1115–27.

10. Savarino V, Di Mario F, Scarpignato C. Proton pump inhibitors in GORDAn overview of their pharmacology, efficacy and safety. *Pharmacol Res* 2009;**59**:135–53.

11. Eusebi LH, Rabitti S, Artesiani ML *et al.* Proton pump inhibitors: Risks of long-term use. *J Gastroenterol Hepatol* 2017;**32**:1295–302.

12. Malfertheiner P, Kandulski A, Venerito M. Proton-pump inhibitors: understanding the complications and risks. *Nat Rev Gastroenterol Hepatol* 2017;**14**:697–710.

13. Chinzon D, Domingues G, Tosetto N *et al.* Saftey of long-term proton pump inhibtiors: facts and myths. *Arq Gastroenterol* 2022;**59**:219–25.

14. Haenisch B, von Holt K, Wiese B *et al.* Risk of dementia in elderly patients with the use of proton pump inhibitors. *Eur Arch Psychiatry Clin Neurosci* 2015;**265**:419–28.

15. Gomm W, von Holt K, Thomé F *et al.* Association of Proton Pump Inhibitors With Risk of Dementia. *JAMA Neurol* 2016;**73**:410.

16. Chen L-Y, Lin H-J, Wu W-T *et al.* Clinical Use of Acid Suppressants and Risk of Dementia in the Elderly: A Pharmaco-Epidemiological Cohort Study. *Int J Environ Res Public Health* 2020;**17**:8271.

17. Torres-Bondia F, Dakterzada F, Galván L *et al.* Proton pump inhibitors and the risk of Alzheimer's disease and non-Alzheimer's dementias. *Sci Rep* 2020;**10**:21046.

18. Lin H-C, Huang K-T, Lin H-L *et al.* Use of gastric acid–suppressive agents increases the risk of dementia in patients with upper gastrointestinal disease: A population-based retrospective cohort study. Forloni G (ed.). *PLoS One* 2021;**16**:e0249050.

19. Zhang Y, Liang M, Sun C *et al.* Proton pump inhibitors use and dementia risk: a meta-analysis of cohort studies. *Eur J Clin Pharmacol* 2020;**76**:139–47.

20. Northuis C, Bell E, Lutsey P *et al.* Cumulative Use of Proton Pump Inhibitors and Risk of Dementia: The Atherosclerosis Risk in Communities Study. *Neurology* 2023, DOI: 10.1212/WNL.0000000000207747.

21. Khan MA, Yuan Y, Iqbal U *et al.* No Association Linking Short-Term Proton Pump Inhibitor Use to Dementia: Systematic Review and Meta-analysis of Observational Studies. *Am J Gastroenterol* 2020;**115**:671–8.

22. Hussain S, Singh A, Zameer S *et al.* No association between proton pump inhibitor use and risk of dementia: Evidence from a meta-analysis. *J Gastroenterol Hepatol* 2020;**35**:19–28.

23. Wu C-L, Lei W-Y, Wang J-S *et al.* Acid suppressants use and the risk of dementia: A population-based propensity score-matched cohort study. Kou YR (ed.). *PLoS One* 2020;**15**:e0242975.

24. Taipale H, Tolppanen A-M, Tiihonen M *et al.* No Association Between Proton Pump Inhibitor Use and

Risk of Alzheimer's Disease. *Am J Gastroenterol* 2017;**112**:1802–8.

25. Gray SL, Walker RL, Dublin S *et al.* Proton Pump Inhibitor Use and Dementia Risk: Prospective Population-Based Study. *J Am Geriatr Soc* 2018;**66**:247–53.

26. Booker A, Jacob LE, Rapp M *et al.* Risk factors for dementia diagnosis in German primary care practices. *Int Psychogeriatrics* 2016;**28**:1059–65.

27. Goldstein FC, Steenland K, Zhao L *et al.* Proton Pump Inhibitors and Risk of Mild Cognitive Impairment and Dementia. *J Am Geriatr Soc* 2017;**65**:1969–74.

28. Cooksey R, Kennedy J, Dennis MS *et al.* Proton pump inhibitors and dementia risk: Evidence from a cohort study using linked routinely collected national health data in Wales, UK. Ramagopalan S V. (ed.). *PLoS One* 2020;**15**:e0237676.

29. Tamim H, Monfared AAT, LeLorier J. Application of lag-time into exposure definitions to control for protopathic bias. *Pharmacoepidemiol Drug Saf* 2007;**16**:250–8.

30. Wilson RS, Leurgans SE, Boyle PA *et al.* Cognitive Decline in Prodromal Alzheimer Disease and Mild Cognitive Impairment. *Arch Neurol* 2011;**68**:1–7.

31. Akter S, Hassan MR, Shahriar M *et al.* Cognitive impact after short-term exposure to different proton pump inhibitors: assessment using CANTAB software. *Alzheimers Res Ther* 2015;**7**:79.

32. Lochhead P, Hagan K, Joshi AD *et al.* Association Between Proton Pump Inhibitor Use and Cognitive Function in Women. *Gastroenterology* 2017;**153**:971-979.e4.

33. Wod M, Hallas J, Andersen K *et al.* Lack of Association Between Proton Pump Inhibitor Use and Cognitive Decline. *Clin Gastroenterol Hepatol* 2018;**16**:681–9.

34. Ahn N, Frenzel S, Wittfeld K *et al.* Lack of association between proton pump inhibitor use and brain aging: a cross-sectional study. *Eur J Clin Pharmacol* 2021;**77**:1039–48.

35. Badiola N, Alcalde V, Pujol A *et al.* The Proton-Pump Inhibitor Lansoprazole Enhances Amyloid Beta Production. Ginsberg SD (ed.). *PLoS One* 2013;**8**:e58837.

36. Fallahzadeh MK, Borhani Haghighi A, Namazi MR. Proton pump inhibitors: predisposers to Alzheimer disease? *J Clin Pharm Ther* 2010;**35**:125–6.

37. Kumar R, Kumar A, Nordberg A *et al.* Proton pump inhibitors act with unprecedented potencies as inhibitors of the acetylcholine biosynthesizing enzyme—A plausible missing link for their association with incidence of dementia. *Alzheimer's Dement* 2020;**16**:1031–42.

38. Lam JR, Schneider JL, Zhao W *et al.* Proton Pump Inhibitor and Histamine 2 Receptor Antagonist Use and Vitamin B 12 Deficiency. *JAMA* 2013;**310**:2435.

39. Moore E, Mander A, Ames D *et al.* Cognitive impairment and vitamin B12: a review. *Int Psychogeriatrics* 2012;**24**:541–56.

40. Luo C, Li M, Qin R *et al.* White Matter Microstructural Damage as an Early Sign of Subjective Cognitive Decline. *Front Aging Neurosci* 2020;**11**:1–11.

41. Mueller ST, Piper BJ. The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *J Neurosci Methods* 2014;**222**:250–9.

42. Corsi PM. Memory and the Medial Temporal Region of the Brain. McGill University, 1972.

43. Baddeley AD, Hitch G. Working Memory. *Current Biology*. Vol 20. 1974, 47–89.

44. Rey A. Mémorisation d'une série de 15 mots en 5 répetitions. In: Presses Universitaries de France (ed.). *L'examen Clinique En Psychologie*. Paris, 1970, 141–193.

45. Fischl B, Salat DH, Busa E *et al.* Whole Brain Segmentation. *Neuron* 2002;**33**:341–55.

46. Fischl B. FreeSurfer. *Neuroimage* 2012;**62**:774–81.

47. MacLullich AMJ, Ferguson KJ, Reid LM *et al.* Higher Systolic Blood Pressure Is Associated With Increased Water Diffusivity in Normal-Appearing White Matter. *Stroke* 2009;**40**:3869–71.

48. Beaulieu C. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed* 2002;**15**:435–55.

49. Wassenaar TM, Yaffe K, van der Werf YD *et al.* Associations between modifiable risk factors and white matter of the aging brain: insights from diffusion tensor imaging studies. *Neurobiol Aging* 2019;**80**:56–70.

50. Cremers LGM, de Groot M, Hofman A *et al.* Altered tract-specific white matter microstructure is related to poorer cognitive performance: The Rotterdam Study. *Neurobiol Aging* 2016;**39**:108–17.

51. Mielke MM, Okonkwo OC, Oishi K *et al.* Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimer's Dement* 2012;**8**:105–13.

52. Kiesel EK, Hopf YM, Drey M. An anticholinergic burden score for German prescribers: score development. *BMC Geriatr* 2018;**18**:239.

53. Cohen S, Williamson GM. Perceived Stress in a Probability Sample of the United States. *Soc Psychol Heal* 1988:31–67.

54. Rubin M. When to adjust alpha during multiple testing: a consideration of disjunction, conjunction, and individual testing. *Synthese* 2021;**199**:10969–1000.

55. Mangoni AA, Jackson SHD. Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. *Br J Clin Pharmacol* 2003;**57**:6–14.

56. Sexton CE, Kalu UG, Filippini N *et al.* A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging* 2011;**32**:2322.e5-2322.e18.

57. Wang Y, West JD, Flashman LA *et al.* Selective changes in white matter integrity in MCI and older adults with cognitive complaints. *Biochim Biophys Acta - Mol Basis Dis* 2012;**1822**:423–30.

58. Power MC, Su D, Wu A *et al.* Association of white matter microstructural integrity with cognition and dementia. *Neurobiol Aging* 2019;**83**:63–72.

59. Puranen A, Taipale H, Koponen M *et al.* Incidence of antidepressant use in community-dwelling persons with and without Alzheimer's disease: 13-year follow-up. *Int J Geriatr Psychiatry* 2017;**32**:94–

101.

60. Esterita T, Dewi S, Suryatenggara FG *et al.* Association of Functional Dyspepsia with Depression and Anxiety: A Systematic Review. *J Gastrointest Liver Dis* 2021;**30**:259–66.

61. Alaeddin N, Stingl JC, Breteler MMB *et al.* Validation of self-reported medication use applying untargeted mass spectrometry-based metabolomics techniques in the Rhineland study. *Br J Clin Pharmacol* 2022;**88**:2380–95.

Table 1 Characteristics of the study population

			Participants with cognition data (n=7,381)				Participants with MRI data (n=4,974)			
	All (n=7,480)	**Missing**	**PPI users (n=413)**	**PPI non-users (n=6,968)**	***p***	***Age adjusted p***	**PPI users (n=250)**	**PPI non-users (n=4,724)**	***p***	***Age adjusted p***
Age (years), M (SD)	55.4 (13.7)	0.0%	64.8 (11.6)	54.6 (13.5)	**<.001**	**---**	64.4 (11.5)	54.3 (13.5)	**<.001**	**---**
Sex (women), N (%)	4,235 (56.6)	0.0%	224 (54.2)	3,953 (56.7)	.346	.748	138 (55.2)	2,760 (58.4)	.346	.280
Education, N (%)		0.0%			**<.001**				**<.001**	
Low	138 (1.8)		160 (38.7)	3,827 (54.9)		**.033**	105 (42.0)	2,653 (56.2)		.161
Middle	3,316 (44.3)		20 (4.8)	110 (1.6)		Ref.	11 (4.4)	67 (1.4)		Ref.
High	4,026 (53.8)		233 (56.4)	3,031 (43.5)		**<.001**	134 (53.6)	2,004 (42.4)		**.009**
Smoking, N (%)		3.5%			**.004**				.065	
Never	3,430 (47.5)		158 (39.6)	3,231 (48.0)		Ref.	102 (41.8)	2,273 (49.0)		Ref.
Former	2,925 (40.5)		189 (47.4)	2,702 (40.1)		.119	112 (45.9)	1,799 (38.8)		.339
Current	861 (11.9)		52 (13.0)	802 (11.9)		**.020**	30 (12.3)	569 (12.3)		.293
BMI (kg/m^2), M (SD)	26.0 (4.6)	1.7%	28.1 (4.8)	25.8 (4.6)	**<.001**	**<.001**	27.5 (4.4)	25.5 (4.1)	**<.001**	**<.001**
Hypertension, N (%)	2,678 (35.9)	0.4%	293 (71.8)	2,314 (33.3)	**<.001**	**<.001**	164 (66.1)	1,509 (32.0)	**<.001**	**<.001**
Diabetes, N (%)	380 (5.3)	4.5%	52 (13.1)	314 (4.7)	**<.001**	**<.001**	26 (10.7)	188 (4.1)	**<.001**	**.017**
Medication use, N (%)										
Statins	767 (10.3)	0.2%	121 (29.7)	611 (8.8)	**<.001**	**<.001**	67 (26.9)	407 (8.6)	**<.001**	**<.001**
Anticholinergics	173 (2.3)	0.5%	33 (8.2)	135 (1.9)	**<.001**	**<.001**	21 (8.6)	88 (1.9)	**<.001**	**<.001**
Antidepressants	458 (6.1)	0.1%	63 (15.5)	384 (5.5)	**<.001**	**<.001**	48 (19.4)	267 (5.7)	**<.001**	**<.001**
NSAID	82 (1.1)	0.3%	38 (9.4)	43 (0.6)	**<.001**	**<.001**	26 (10.5)	23 (0.5)	**<.001**	**<.001**
Antithrombotic	785 (10.5)	0.1%	130 (31.9)	615 (8.8)	**<.001**	**<.001**	71 (28.5)	387 (8.2)	**<.001**	**<.001**
Cognitive domain scores (z-standardized)										
Global cognition	0.0 (0.6)	1.3%	-0.4 (0.6)	0.0 (0.6)	**<.001**	**<.001**	-0.3 (0.6)	0.1 (0.6)	**<.001**	**.005**
Total memory	0.0 (0.7)	0.8%	-0.4 (0.7)	0.0 (0.7)	**<.001**	**<.001**	-0.3 (0.7)	0.1 (0.7)	**<.001**	**.024**
Executive function	0.0 (0.8)	0.8%	-0.4 (0.8)	0.0 (0.7)	**<.001**	**.002**	-0.4 (0.8)	0.0 (0.7)	**<.001**	**.027**
Processing speed	0.0 (0.8)	0.4%	-0.4 (0.9)	0.1 (0.8)	**<.001**	.061	-0.4 (0.9)	0.1 (0.8)	**<.001**	.103
Working memory	0.0 (0.7)	0.6%	-0.3 (0.7)	0.0 (0.7)	**<.001**	**.001**	-0.3 (0.7)	0.0 (0.7)	**<.001**	**.015**
Episodic verbal memory	0.0 (0.9)	0.3%	-0.4 (1.0)	0.1 (0.9)	**<.001**	**.004**	-0.3 (1.0)	0.1 (0.9)	**<.001**	.154
Macrostructural brain measures (ml)										
Total brain volume	1105.8 (117.6)	29.1%	1062.8 (109.6)	1109.0 (117.2)	**<.001**	.207	1058.4 (106.6)	1107.7 (116.1)	**<.001**	.107
Cortical grey matter volume	459.0 (48.0)	29.0%	439.2 (43.6)	460.5 (47.8)	**<.001**	**.034**	437.4 (42.2)	459.9 (47.4)	**<.001**	**.015**
White matter volume	455.94 (58.77)	28.9	438.6 (56.0)	457.3 (58.7)	**<.001**	.475	436.9 (54.4)	456.7 (57.9)	**<.001**	.324
Ventricle volume	28.7 (15.0)	29.6%	34.9 (17.4)	28.3 (14.7)	**<.001**	.969	34.4 (17.2)	28.2 (14.5)	**<.001**	.772
Hippocampal volume (left hemisphere)	3.8 (0.4)	28.8%	3.8 (0.4)	3.9 (0.4)	**<.001**	.530	3.7 (0.4)	3.9 (0.4)	**<.001**	.736
Hippocampal volume (right hemisphere)	4.0 (0.48)	28.8%	3.8 (0.5)	4.0 (0.5)	**<.001**	.639	3.8 (0.5)	4.0 (0.5)	**<.001**	.586
Cortical thickness (left hemisphere)	2.4 (0.1)	29.0%	2.4 (0.1)	2.4 (0.1)	**<.001**	.581	2.4 (0.1)	2.4 (0.1)	**<.001**	.574
Cortical thickness (right hemisphere)	2.4 (0.1)	29.0%	2.4 (0.1)	2.4 (0.1)	**<.001**	.869	2.4 (0.1)	2.4 (0.1)	**<.001**	.733
Microstructural brain measures										
Global FA	0.6 (0.0)	32.6%	0.6 (0.0)	0.6 (0.0)	**<.001**	**.037**	0.6 (0.0)	0.6 (0.0)	**<.001**	**.020**
Global MD (10^-4 mm^2/s)	0.0 (0.0)	32.6%	0.0 (0.0)	0.0 (0.0)	**<.001**	.105	0.0 (0.0)	0.0 (0.0)	**<.001**	.087

Abbreviations: BMI, Body Mass Index; FA, fractional anisotropy; M, mean; MD, mean diffusivity; n, number of participants; NSAID, non-steroidal anti-inflammatory drugs; Ref, reference; SD, standard deviation.
Note. Data presented as mean ± standard deviation for quantitative variables/ percentages for categorical variables. *p*-values in bold represent significant values. Differences between PPI users and non-users were assessed with logistic regression adjusted for age.

Table 2 Linear regression β coefficients (95%CI) for PPI use on z-standardized cognitive domain scores (stratified by age groups and duration of use)

		All (n=7,381; PPI users: n=413)			<65 years (n=5,482; PPI users: n=199)			≥65 years (n=1,899; PPI users: n=214)		
		β	***95% CI***	***p***	**β**	***95% CI***	***p***	**β**	***95% CI***	***p***
All PPI users	***Global cognition***	-0.04	-0.08 – 0.01	.104	**-0.08**	**-0.14 – -0.01**	**.020**	-0.01	-0.08 – 0.06	.880
	Total memory	-0.04	-0.10 – 0.01	.139	**-0.09**	**-0.17 – -0.01**	**.029**	-0.00	-0.09 – 0.08	.926
	Executive function*	-0.03	-0.10 – 0.03	.332	-0.07	-0.16 – 0.02	.105	0.01	-0.09 – 0.12	.812
	Processing speed	-0.02	-0.08 – 0.05	.635	-0.02	-0.11 – 0.07	.692	-0.02	-0.13 – 0.09	.710
	Working memory	-0.05	-0.12 – 0.02	.147	**-0.13**	**-0.22 – -0.03**	**.009**	0.03	-0.07 – 0.12	.589
	Episodic verbal memory*	-0.04	-0.12 – 0.05	.409	-0.05	-0.16 – 0.07	.413	-0.03	-0.16 – 0.10	.616
Short-term users (<3 years)		**All Short-term users (n=164)**			**Short-term users <65 years (n=94)**			**Short-term users ≥65 years (n=70)**		
		β	***95% CI***	***p***	**β**	***95% CI***	***p***	**β**	***95% CI***	***p***
	Global cognition	-0.02	-0.09 – 0.05	.575	-0.07	-0.16 – 0.02	.131	0.04	-0.07 – 0.16	.463
	Total memory	-0.01	-0.11 – 0.08	.770	-0.06	-0.18 – 0.06	.344	0.04	-0.10 – 0.19	.549
	Executive function*	0.01	-0.09 – 0.11	.829	-0.07	-0.20 – 0.06	.267	0.12	-0.05 – 0.29	.181
	Processing speed	-0.04	-0.14 – 0.07	.493	-0.06	-0.19 – 0.07	.348	-0.01	-0.19 – 0.17	.910
	Working memory	-0.03	-0.14 – 0.07	.528	-0.12	-0.25 – 0.02	.101	0.07	-0.08 – 0.23	.363
	Episodic verbal memory*	0.02	-0.11 – 0.14	.817	0.01	-0.15 – 0.18	.895	0.02	-0.20 – 0.24	.847
Long-term users (≥3 years)		**All Long-term users (n=248)**			**Long-term users <65 years (n=105)**			**Long-term users ≥65 years (n=143)**		
		β	***95% CI***	***p***	**β**	***95% CI***	***p***	**β**	***95% CI***	***p***
	Global cognition	-0.05	-0.11 – 0.01	.092	-0.08	-0.16 – 0.00	.065	-0.03	-0.11 – 0.06	.517
	Total memory	-0.06	-0.14 – 0.01	.090	**-0.12**	**-0.23 – -0.01**	**.030**	-0.03	-0.13 – 0.08	.632
	Executive function*	-0.06	-0.14 – 0.02	.143	-0.07	-0.19 – 0.04	.216	-0.04	-0.16 – 0.08	.540
	Processing speed	-0.00	-0.09 – 0.08	.950	0.02	-0.10 – 0.14	.766	-0.02	-0.15 – 0.11	.726
	Working memory	-0.06	-0.15 – 0.02	.139	**-0.14**	**-0.26 – -0.01**	**.032**	-0.00	-0.11 – 0.11	.994
	Episodic verbal memory*	-0.06	-0.17 – 0.04	.239	-0.09	-0.24 – 0.05	.212	-0.05	-0.20 – 0.10	.516

Models adjusted for age, age^2, sex, education, smoking, German mother tongue (*only for executive function and episodic verbal memory), BMI, hypertension, diabetes, antithrombotic medication, antidepressants, statins, anticholinergic medication, NSAIDs, perceived stress.
Abbreviations: BMI, body mass index; CI, confidence interval; n, number of participants; NSAID, non-steroidal anti-inflammatory drugs; PPI, proton pump inhibitors;
Note. Values in bold represent significant values.

Table 3 Linear regression β coefficients (95%CI) for PPI use on z-standardized macrostructural brain measures (stratified by age groups and duration of use)

	All (n=4,974; 250 PPI users)			**Short-term PPI users (n=102)**			**Long-term PPI users (n=148)**		
	β	***95% CI***	***p***	**β**	***95% CI***	***p***	**β**	***95% CI***	***p***
Total brain volume	0.01	-0.03 – 0.06	.635	-0.03	-0.10 – 0.04	.401	0.04	-0.02 – 0.10	.189
Cortical grey matter volume	-0.01	-0.07 – 0.04	.629	-0.05	-0.14 – 0.04	.249	0.01	-0.06 – 0.08	.768
White matter volume	0.02	-0.05 – 0.08	.603	-0.02	-0.12 – 0.07	.623	0.04	-0.04 – 0.12	.276
Ventricle volume	0.01	-0.10 – 0.11	.895	0.08	-0.08 – 0.24	.322	-0.04	-0.17 – 0.09	.530
Hippocampal volume (left hemisphere)	-0.01	-0.11 – 0.09	.867	0.03	-0.12 – 0.18	.704	-0.03	-0.16 – 0.09	.595
Hippocampal volume (right hemisphere)	-0.01	-0.11 – 0.09	.874	-0.01	-0.16 – 0.14	.883	-0.01	-0.13 – 0.12	.928
Cortical thickness (left hemisphere)	0.02	-0.10 – 0.13	.788	-0.02	-0.20 – 0.16	.852	0.04	-0.11 – 0.19	.612
Cortical thickness (right hemisphere)	0.02	-0.10 – 0.15	.712	-0.02	-0.21 – 0.17	.839	0.05	-0.10 – 0.21	.513

Models adjusted for age, age^2, sex, education, smoking, BMI, hypertension, diabetes, antithrombotic medication, antidepressants, statins, anticholinergic medication, NSAIDs, perceived stress, estimated intracranial volume.
Abbreviations: BMI, body mass index; CI, confidence interval; n, number of participants; NSAID, non-steroidal anti-inflammatory drugs; PPI, proton pump inhibitors;
Note. Values in bold represent significant values.

Table 4 Linear regression β coefficients (95%CI) for PPI use on z-standardized microstructural brain measures (stratified by age groups and duration of use)

	All (n=4,974; 250 PPI users)			Short-term PPI users (n=102)			Long-term PPI users (n=148)		
	β	***95% CI***	***p***	**β**	***95% CI***	***p***	**β**	***95% CI***	***p***
Global FA	-0.05	-0.18 – 0.07	.400	-0.12	-0.31 – 0.07	.206	-0.01	-0.17 – 0.15	.926
Global MD	**0.11**	**0.00 – 0.23**	**.048**	0.13	-0.04 – 0.30	.131	0.10	-0.04 – 0.24	.161
Body of corpus callosum (MD)	**0.20**	**0.08 – 0.33**	**.002**	**0.20**	**0.01 – 0.39**	**.037**	**0.20**	**0.04 – 0.36**	**.014**
Left cingulum (cingulate gyrus)	**0.13**	**0.00 – 0.25**	**.048**	0.15	-0.04 – 0.34	.126	0.11	-0.05 – 0.27	.166
Right cingulum (cingulate gyrus) (MD)	0.08	-0.05 – 0.21	.233	0.08	-0.11 – 0.27	.425	0.08	-0.08 – 0.24	.349
Left cingulum (hippocampus) (MD)	0.05	-0.08 – 0.18	.459	**0.21**	**0.01 – 0.41**	**.042**	-0.06	-0.23 – 0.11	.504
Right cingulum (hippocampus) (MD)	-0.05	-0.20 – 0.09	.450	0.15	-0.07 – 0.36	.184	-0.19	-0.37 – -0.01	**.067**
Left corticospinal tract (MD)	0.05	-0.09 – 0.19	.520	-0.12	-0.33 – 0.10	.284	0.15	-0.02 – 0.33	.084
Right corticospinal tract (MD)	0.12	-0.02 – 0.26	.098	0.01	-0.20 – 0.23	.904	**0.19**	**0.01 – 0.37**	**.036**
Fornix (column and body) (MD)	0.00	-0.13 – 0.13	.993	0.01	-0.19 – 0.21	.939	-0.00	-0.17 – 0.16	.959
Genu of corpus callosum (MD)	0.03	-0.10 – 0.16	.664	0.04	-0.17 – 0.24	.728	0.03	-0.14 – 0.19	.768
Left posterior thalamic radiation (MD)	-0.00	-0.12 – 0.12	.959	0.05	-0.14 – 0.23	.613	-0.04	-0.19 – 0.12	.632
Right posterior thalamic radiation (MD)	0.02	-0.10 – 0.15	.716	0.09	-0.10 – 0.29	.362	-0.02	-0.18 – 0.14	.796
Splenium of corpus callosum (MD)	0.12	-0.01 – 0.26	.070	0.06	-0.14 – 0.26	.565	0.17	-0.00 – 0.33	.052
Left superior longitudinal fasciculus (MD)	0.09	-0.03 – 0.21	.129	0.10	-0.08 – 0.28	.269	0.08	-0.06 – 0.23	.264
Right superior longitudinal fasciculus (MD)	0.12	-0.00 – 0.24	.056	0.14	-0.05 – 0.32	.147	0.11	-0.05 – 0.26	.174
Left sagittal stratum (MD)	-0.01	-0.13 – 0.11	.878	-0.03	-0.21 – 0.16	.782	0.00	-0.15 – 0.15	.983
Right sagittal stratum (MD)	0.06	-0.06 – 0.19	.331	0.05	-0.15 – 0.24	.628	0.07	-0.09 – 0.23	.370
Left uncinate fasciculus (MD)	0.07	-0.07 – 0.20	.325	**0.22**	**0.02 – 0.43**	**.035**	-0.04	-0.21 – 0.14	.684
Right uncinate fasciculus (MD)	0.11	-0.02 – 0.25	.108	0.12	-0.09 – 0.32	.271	0.11	-0.06 – 0.28	.215

Models adjusted for age, age^2, sex, education, smoking, BMI, hypertension, diabetes, antithrombotic medication, antidepressants, statins, anticholinergic medication, NSAIDs, perceived stress.
Abbreviations: BMI, body mass index; CI, confidence interval; FA, fractional anisotropy; MD, mean diffusivity; n, number of participants; NSAID, non-steroidal anti-inflammatory drugs; PPI, proton pump inhibitors;
Note. Values in bold represent significant values.

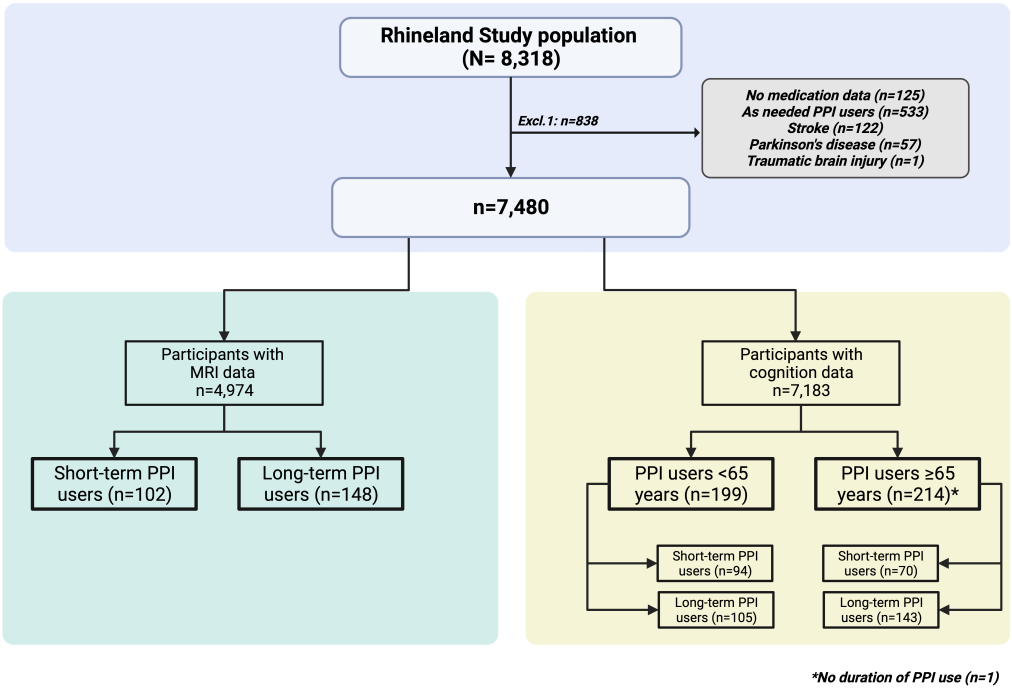


Figure 1 **Sample flow chart** (created with BioRender)

APPENDIX A: Cognitive test battery

Verbal Learning and Memory Test

We used the German version of the 15-word Verbal Learning and Memory Test (VLMT), which is analogous to the Rey Auditory Verbal Learning Test [1,2]. Here, 15 semantically unrelated nouns are learned and recalled over multiple trials, and declarative episodic verbal memory (short- and long-term memory) and learning performance are measured. The test begins with five trials of auditory learning and recall, followed by recall of an interference list, another immediate recall and delayed recall after 20-30 minutes. The outcome measures were the number of words recalled correctly in the immediate recall (sum of recalls one to five) and the number of words recalled correctly after the time delay.

Digit Span Task

The Digit Span Task is a verbal working memory task in which participants were asked to recall sequences of digits of increasing length in forward (sequence length 3-9) and backward order (sequence length 2-9). After two errors within a sequence in the forward and backward tests, the task ends. The maximum sequence length was used as the outcome.

Corsi Block-tapping Test

The Corsi Block-tapping Test is a verbal working memory test [3]. Participants are asked to recall visuospatial sequences of blocks that changed colour by tapping the blocks in the correct order on a tablet. In the forward task (sequence length 2-9), participants tap the blocks in forward order, and in the backward task (sequence length 2-9), participants tap the blocks in backward order. After two errors within a sequence, the task ended and the length of the last sequence successfully completed was used as the result.

Trail-making Test

The Trail-making Test (TMT) assesses processing speed and executive function and has been adapted from the Psychology Experiment Building Language (PEBL) test battery [4] to a touch screen. In version A of the TMT, numbers from 1 to 25 are randomly scattered on the screen and participants need to connect them in ascending order (1-2-3- etc) as quickly as possible. In version B, 12 numbers (1 to 12) and 12 letters (A to L) are scattered randomly on the screen and have to be connected in ascending order and in alternation (1-A-2-B etc) as quickly as possible. The main outcome is the time taken to complete the task in both versions. If the participant takes more than 301 seconds, the test is automatically stopped.

Word Fluency Task

The Word Fluency Task assesses semantic memory and executive function by asking participants to name as many animals as possible in one minute. The score is based on the number of animals named correctly.

Pro-saccade and Anti-saccade tasks

The Pro-saccade and Anti-saccade tasks are part of the Eye-tracking examination [5] and are used as a cognitive measure of attention, processing speed and executive function. In both tasks, a 1° diameter white circle is presented on a black background. In the Pro-saccade task, a stimulus appeared at the central position for a randomly determined duration. The stimulus moved to a horizontal side position where it remained for 1,000 ms before moving back to the central position (random order of 15 left-sided and 15 right-sided trials). Participants were instructed to follow the stimulus as closely as possible. Subsequently, the anti-saccade task was performed in the same way, except that participants were instructed to look at the stimulus when it was in the central position and immediately look at the opposite (mirror) position when the stimulus moved to the side. Prosaccade latency (time needed to initiate a saccade) and antisaccade error rate (percentage of trials, in which the participant makes a direction error) were used as outcome measures.

APPENDIX B: MRI acquisition and processing

The 3T MRI scanners (MAGNETOM Prisma, Siemens Healthineers, Erlangen, Germany) are equipped with a 64-channel phased-array head/neck coil. T1-weighted images were acquired with an isotropic spatial resolution of 0.8 mm using a multi-echo MPRAGE sequence (acquisition time = 6.5 min, repetition time = 2560 ms, inversion time = 1100 ms, flip angle 7°, field of view = 256 x 256 mm, 224 sagittal slices). Structural volumes and thicknesses were determined using the standard FreeSurfer processing pipeline (http://surfer.nmr.mgh.harvard.edu/) [6,7] on T1-weighted MR images. We used the estimated total intracranial volume (eTIV) generated by FreeSurfer as a proxy for head size [8].

Simultaneous-multi-slice diffusion weighted MRI (dMRI) was performed using a spin-echo echo-planar imaging (SE-EPI) sequence applying threefold slice-acceleration [6,7,9]. A compressed sensing [10] diffusion spectrum imaging [11] (CS-DSI) protocol [12] was used to acquire dMRI scans with an isotropic spatial resolution of 1.5 mm (acquisition time = 12.1 min, repetition time = 5500 ms, echo time = 105ms, field of view =210x210mm, 96slices, diffusion weighting =6800s/mm2, gradient pulse separation =49.5ms, gradient pulse duration =19.7ms). After correction of susceptibility-induced [13] and eddy-current-induced geometric distortions and subject motion [12,14] using FMRIB Software Library (FSL) version 6.0 (www.fmrib.ox. ac.uk/fsl), CS reconstruction recovered 257 unique DWIs from 112 undersampled DSI acquisitions [12,15]. FA and MD are estimated by voxel-wise model fitting using the Microstructure

Diffusion Toolbox (MDT) [16,17]. A whole brain white matter mask was obtained from the T1-weighted MR image using the standard Freesurfer processing pipeline. This mask was further corrected for white matter hyperintensities obtained from T1-weighted, T2-weighted and FLAIR images and further refined through FA skeletonization. Applying this mask, global dMRI measures were computed as the average across voxels within normal appearing white matter (WM). Additionally, WM tract-specific dMRI measures were derived for the regions of interest, that have been associated with cognition [18,19], provided by the JHU-ICBM DTI atlas [20].

References

1. Baddeley AD, Hitch G. Working Memory. *Current Biology*. Vol 20. 1974, 47–89.

2. Rey A. Mémorisation d'une série de 15 mots en 5 répetitions. In: Presses Universitaries de France (ed.). *L'examen Clinique En Psychologie*. Paris, 1970, 141–193.

3. Corsi PM. Memory and the Medial Temporal Region of the Brain. McGill University, 1972.

4. Mueller ST, Piper BJ. The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *J Neurosci Methods* 2014;**222**:250–9.

5. Coors A, Merten N, Ward DD *et al.* Strong age but weak sex effects in eye movement performance in the general adult population: Evidence from the Rhineland Study. *Vision Res* 2021;**178**:124–33.

6. Fischl B, Salat DH, Busa E *et al.* Whole Brain Segmentation. *Neuron* 2002;**33**:341–55.

7. Fischl B. FreeSurfer. *Neuroimage* 2012;**62**:774–81.

8. Buckner RL, Head D, Parker J *et al.* A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage* 2004;**23**:724–38.

9. Schirmer MD, Dalca A V., Sridharan R *et al.* White matter hyperintensity quantification in large-scale clinical acute ischemic stroke cohorts – The MRI-GENIE study. *NeuroImage Clin* 2019;**23**:101884.

10. Setsompop K, Gagoski BA, Polimeni JR *et al.* Blipped-Controlled Aliasing in Parallel Imaging (blipped-CAIPI) for simultaneous multi-slice EPI with reduced g-factor penalty. *Magn Reson Med* 2012;**67**:1210–24.

11. Xu J, Moeller S, Auerbach EJ *et al.* Evaluation of slice accelerations using multiband echo planar imaging at 3T. *Neuroimage* 2013;**83**:991–1001.

12. Cauley SF, Polimeni JR, Bhat H *et al.* Interslice leakage artifact reduction technique for simultaneous multislice acquisitions. *Magn Reson Med* 2014;**72**:93–102.

13. Menzel MI, Tan ET, Khare K *et al.* Accelerated diffusion spectrum imaging in the human brain using compressed sensing. *Magn Reson Med* 2011;**66**:1226–33.

14. Wedeen VJ, Hagmann P, Tseng WI *et al.* Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn Reson Med* 2005;**54**:1377–86.

15. Tobisch A, Stirnberg R, Harms RL *et al.* Compressed Sensing Diffusion Spectrum Imaging for Accelerated Diffusion Microstructure MRI in Long-Term Population Imaging. *Front Neurosci* 2018;**12**:1–19.

16. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 2003;**20**:870–88.

17. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 2016;**125**:1063–78.

18. Cremers LGM, de Groot M, Hofman A *et al.* Altered tract-specific white matter microstructure is related to poorer cognitive performance: The Rotterdam Study. *Neurobiol Aging* 2016;**39**:108–17.

19. Mielke MM, Okonkwo OC, Oishi K *et al.* Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimer's Dement* 2012;**8**:105–13.

20. Bürgel U, Amunts K, Hoemke L *et al.* White matter fiber tracts of the human brain: Three-dimensional mapping at microscopic resolution, topography and intersubject variability. *Neuroimage* 2006;**29**:1092–105.

4 Discussion

This general discussion begins with a brief summary of the key findings, implications, methodological considerations, and future perspectives for each chapter. Thereafter, in an overarching conclusion I will tie together the central themes and contributions of this thesis.

4.1 Quality of medication data

In **Chapter 3.1**, advanced mass spectrometry-based metabolomics were used to validate self-reported medication data. The results revealed high concordances between self-reported medication data and metabolite measurements, independent of age and sex.

This not only minimizes concerns about recall and reporting bias (Hafferty *et al.*, 2018), but also bears promising implications for leveraging population-based cohort data in epidemiological research. Our assessment of a broad spectrum of commonly used drugs and their metabolites provides information across various drug classes. Importantly, this contribution adds to the discourse on self-reported medication data, particularly in light of conflicting results observed in prior studies during validation with EMRs (Noize *et al.*, 2009; Hafferty *et al.*, 2018).

While our research provides valuable insights, it comes with certain limitations. We excluded drugs with shorter half-lives, as the primary focus was on medications with sufficiently long half-lives. This decision was made to avoid low concordance rates, as drugs with shorter half-lives might not always be detectable in blood. Another limitation of our study is that the accuracy of self-reported medication use may be culture-dependent, particularly for drugs with indications susceptible to stigma.

Our research paves the way for future endeavors, including the combination of various data sources, such as primary and secondary data (March, 2017). Combining these data would allow for comprehensive data on medication use and related health conditions. Looking ahead, a synergistic approach combining self-reported data, EMRs, and blood-based metabolite measurements would provide the highest-quality information on drug intake. Population-based cohort studies, like the Rhineland Study, are well-suited for this purpose, as they offer real-world data on medication use instead of a controlled setting. Combining diverse data sources with deep-phenotyping techniques, such as brain

imaging and disease biomarkers, ensures a holistic understanding of medication impact on health.

4.2 Understanding and addressing over- and undertreatment

Chapters 3.2 and **3.3** centered on the prevalence and putative determinants of inadequate treatment with commonly prescribed drugs. **Chapter 3.2** focused on levothyroxine, revealing that nearly a quarter of our study population were regular users. Among users, 18% were overtreated, while 4% were undertreated. Notably, the elderly were at higher risk of overtreatment. **Chapter 3.3** explored the use of antihypertensive medications, finding that over half of users received inadequate treatment (20% overtreated; 33% undertreated). We also observed sex differences, where women were more likely to be overtreated compared to men.

Balancing drug therapy is comparable to a perpetual scale, where tilting too far in either direction presents distinct challenges. Overtreatment can lead to unnecessary side effects and increased healthcare costs (Kojima *et al.*, 2012; Lyu *et al.*, 2017), while undertreatment might prolong suffering, compromise health outcomes, and lead to additional expenditures (Mamede and Schmidt, 2014; Kearney, Treadwell and Marshall, 2017). Therefore, over- and undertreatment are often referred to as the '*conjoined twins of modern medicine*' (Heath, 2014). In conditions like hypothyroidism and hypertension, the '*lower-the-better*' approach is recognized to have potential drawbacks (Biondi and Cooper, 2018; Williams *et al.*, 2018), making both overtreatment and undertreatment equally undesirable and potentially dangerous. We showed that the prevalence of inadequate treatment is high, and identified at-risk populations. Importantly, we found sex-specific differences in the determinants of inadequate treatment for hypertension. These findings should be incorporated into treatment guidelines to improve well-being and quality of life, especially given the current oversight of sex differences.

The generalizability of our findings, particularly for levothyroxine, might be limited to Germany, as our population exhibits a higher use of this drug compared to the European average (Okosieme *et al.*, 2011; Wouters *et al.*, 2020; Janett-Pellegri *et al.*, 2021). This variation is possibly influenced by Germany's focus on thyroid-related medical procedures (Verburg, 2015; Kiel *et al.*, 2020). Moreover, our study primarily included well-educated participants, which enhanced data quality and reliability, but might introduce bias,

potentially leading to an underestimation of suboptimal treatment prevalence. We were also unable to track reasons for suboptimal treatment, such as patient-related factors like non-adherence, or assess efforts to achieve treatment goals.

Future research should expand on the assessment of over- and undertreatment across other drug classes, prioritize personalized and precision medicine-based treatment strategies, and thoroughly explore the underlying causes of treatment responsiveness. This might involve considering genetic, epigenetic, and metabolic factors. Moreover, comprehensive treatment guidelines addressing both over- and undertreatment, with a focus on sex-specific variations in treatment protocols, need to be developed.

4.3 Impact of PPIs on cognition and brain

In **Chapter 3.4,** we investigated the relation between PPIs, cognitive function, and brain structures. Prolonged PPI use was associated with poorer cognitive performance, particularly in younger long-term users. While volumetric brain measures showed no significant associations, microstructural brain measures showed higher mean diffusivity in PPI users compared to non-users.

PPIs are widely used in Europe and the USA (Heidelbaugh *et al.*, 2012; Lanas, 2016; Schumock *et al.*, 2016; Johnson *et al.*, 2017), but up to 40% of prescriptions are considered inappropriate due to drug-drug interactions, prolonged PPI use, and elevated risk-benefit ratios (Heidelbaugh, Goldberg and Inadomi, 2010; Yang and Metz, 2010; Pasina *et al.*, 2011). Research has raised concerns about their safety, and although studies on the PPI-dementia link have produced mixed results (Haenisch *et al.*, 2015; Gomm *et al.*, 2016; Wod *et al.*, 2018; Hussain *et al.*, 2020), our research showed associations between PPI use, cognitive decline, and changes in brain microstructure. This shows the need for further research, a thorough re-evaluation of PPI safety, and a deeper exploration of causal mechanisms to ensure appropriate PPI use.

Our study extends the existing literature as it explores microstructural brain measures for early detection of changes in white matter integrity (Ahn *et al.*, 2021). A particular strength of our study is its comprehensive medication data, which also included over-the-counter PPIs. However, the cross-sectional nature of our data limited our ability to establish temporal associations. Therefore, we cannot confirm whether PPI use preceded the observed changes in cognitive function and brain microstructural measures. Further

research with longitudinal data is indispensable for deeper insights. The Rhineland Study's design and rich data offer promise for further extensive investigations into the long-term consequences of PPI usage.

These findings suggest that future treatment strategies should consider reducing inappropriate PPI use, while also being mindful of possible confounding by indication. Given the high dementia prevalence in Germany (Thyrian *et al.*, 2020), addressing modifiable risk factors, such as unnecessary PPI prescriptions, is crucial. Inappropriate PPI use often stems from ulcer treatment, stress prevention, misdiagnoses of acid-related disorders, and short-term non-steroidal anti-inflammatory drug use (Savarino *et al.*, 2017). Therefore, educational initiatives for both patients and prescribing doctors are essential to promote safe PPI use. Further research should aim to replicate our findings and include longitudinal data to gain comprehensive understanding of modifiable risk factors.

4.4 Conclusion

We have demonstrated the high quality of our self-reported medication data, coupled with our rich data, affirming its suitability for rigorous research. Particularly valuable for drug utilization studies, our data accurately mirrors actual drug intake. Uncovering a significant prevalence of inadequate treatment with commonly used drugs underscores the urgency to optimize treatment strategies. Moreover, we have shown the enhanced value derived from the integration of self-reported medication data with innovative measurements, such as untargeted mass spectrometry-based metabolomics and state-of-the-art MRI measurements, thereby contributing and expanding upon the existing literature.

Our research not only highlights critical issues but also prompts broader discussions on medication use and its unintended impact on overall health. It emphasizes the importance of collaboration among physicians, pharmacists, and patients for more effective treatment strategies. Population-based cohorts enable long-term real-life assessments, providing invaluable insights into drug effects beyond the confines of controlled clinical trials. By fostering collaboration between researchers and healthcare professionals, pharmacoepidemiology paves the way for a future where patients receive safe and effective treatment, ultimately enhancing their well-being. As Paracelsus wisely noted, '*the dose makes the poison*', emphasizing the pivotal role of considerate dosing in achieving therapeutic balance and guiding our pursuit of better healthcare.

4.5 References

Ahn, N. *et al.* (2021) 'Lack of association between proton pump inhibitor use and brain aging: a cross-sectional study', *European Journal of Clinical Pharmacology*. European Journal of Clinical Pharmacology, 77(7), pp. 1039–1048. doi: 10.1007/s00228-020-03068-8.

Biondi, B. and Cooper, D. S. (2018) 'Subclinical Hyperthyroidism.', *The New England journal of medicine*. CRC Press, 378(25), pp. 2411–2419. doi: 10.1056/NEJMcp1709318.

Gomm, W. *et al.* (2016) 'Association of Proton Pump Inhibitors With Risk of Dementia', *JAMA Neurology*, 73(4), p. 410. doi: 10.1001/jamaneurol.2015.4791.

Haenisch, B. *et al.* (2015) 'Risk of dementia in elderly patients with the use of proton pump inhibitors', *European Archives of Psychiatry and Clinical Neuroscience*, 265(5), pp. 419–428. doi: 10.1007/s00406-014-0554-0.

Hafferty, J. D. *et al.* (2018) 'Self-reported medication use validated through record linkage to national prescribing data', *Journal of Clinical Epidemiology*. Elsevier Inc, 94, pp. 132–142. doi: 10.1016/j.jclinepi.2017.10.013.

Heath, I. (2014) 'Role of fear in overdiagnosis and overtreatment--an essay by Iona Heath.', *BMJ (Clinical research ed.)*, 349(aug06 2), p. g6123. doi: 10.1136/bmj.g6123.

Heidelbaugh, J. J. *et al.* (2012) 'Overutilization of proton-pump inhibitors: what the clinician needs to know', *Therapeutic Advances in Gastroenterology*, 5(4), pp. 219–232. doi: 10.1177/1756283X12437358.

Heidelbaugh, J. J., Goldberg, K. L. and Inadomi, J. M. (2010) 'Magnitude and economic effect of overuse of antisecretory therapy in the ambulatory care setting.', *The American journal of managed care*, 16(9), pp. e228-34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21250399.

Hussain, S. *et al.* (2020) 'No association between proton pump inhibitor use and risk of dementia: Evidence from a meta-analysis', *Journal of Gastroenterology and Hepatology*, 35(1), pp. 19–28. doi: 10.1111/jgh.14789.

Janett-Pellegri, C. *et al.* (2021) 'Prevalence and factors associated with chronic use of levothyroxine: A cohort study', *PLOS ONE*. Edited by A. Lupattelli, 16(12), p. e0261160.

doi: 10.1371/journal.pone.0261160.

Johnson, D. A. *et al.* (2017) 'The Safety of Appropriate Use of Over-the-Counter Proton Pump Inhibitors: An Evidence-Based Review and Delphi Consensus', *Drugs*. Springer International Publishing, 77(5), pp. 547–561. doi: 10.1007/s40265-017-0712-6.

Kearney, M., Treadwell, J. and Marshall, M. (2017) 'Overtreatment and undertreatment: time to challenge our thinking', *British Journal of General Practice*, 67(663), pp. 442–443. doi: 10.3399/bjgp17X692657.

Kiel, S. *et al.* (2020) 'Frequency of thyroid function tests and examinations in participants of a population-based study.', *BMC health services research*. BMC Health Services Research, 20(1), p. 70. doi: 10.1186/s12913-020-4910-7.

Kojima, G. *et al.* (2012) 'Reducing Cost by Reducing Polypharmacy: The Polypharmacy Outcomes Project', *Journal of the American Medical Directors Association*, 13(9), pp. 818.e11-818.e15. doi: https://doi.org/10.1016/j.jamda.2012.07.019.

Lanas, A. (2016) 'We Are Using Too Many PPIs, and We Need to Stop: A European Perspective', *American Journal of Gastroenterology*. Nature Publishing Group, 111(8), pp. 1085–1086. doi: 10.1038/ajg.2016.166.

Lyu, H. *et al.* (2017) 'Overtreatment in the United States', *PLoS ONE*, 12(9), pp. 1–11. doi: 10.1371/journal.pone.0181970.

Mamede, S. and Schmidt, H. G. (2014) 'The twin traps of overtreatment and therapeutic nihilism in clinical practice.', *Medical education*, 48(1), pp. 34–43. doi: 10.1111/medu.12264.

March, S. (2017) 'Individual Data Linkage of Survey Data with Claims Data in Germany—An Overview Based on a Cohort Study', *International Journal of Environmental Research and Public Health*, 14(12), p. 1543. doi: 10.3390/ijerph14121543.

Noize, P. *et al.* (2009) 'Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study', *Pharmacoepidemiology and Drug Safety*, 18(4), pp. 310–319. doi: 10.1002/pds.1717.

Okosieme, O. E. *et al.* (2011) 'Adequacy of thyroid hormone replacement in a general population', *Qjm*, 104(5), pp. 395–401. doi: 10.1093/qjmed/hcq222.

Pasina, L. *et al.* (2011) ‘Prevalence and appropriateness of drug prescriptions for peptic ulcer and gastro-esophageal reflux disease in a cohort of hospitalized elderly’, *European Journal of Internal Medicine*. European Federation of Internal Medicine., 22(2), pp. 205–210. doi: 10.1016/j.ejim.2010.11.009.

Savarino, V. *et al.* (2017) ‘The appropriate use of proton pump inhibitors (PPIs): Need for a reappraisal’, *European Journal of Internal Medicine*. European Federation of Internal Medicine., 37, pp. 19–24. doi: 10.1016/j.ejim.2016.10.007.

Schumock, G. T. *et al.* (2016) ‘National trends in prescription drug expenditures and projections for 2016’, *American Journal of Health-System Pharmacy*, 73(14), pp. 1058–1075. doi: 10.2146/ajhp160205.

Thyrian, J. R. *et al.* (2020) ‘Die Prävalenz an Demenz erkrankter Menschen in Deutschland – eine bundesweite Analyse auf Kreisebene’, *Der Nervenarzt*, 91(11), pp. 1058–1061. doi: 10.1007/s00115-020-00923-y.

Verburg, F. A. (2015) ‘Is thyroid surgery performed too often in Germany?’, *Nuklearmedizin*, 54(03), pp. 101–105. doi: 10.1055/s-0037-1616609.

Williams, B. *et al.* (2018) ‘2018 ESC/ESH Guidelines for the management of arterial hypertension’, *European Heart Journal*, 39(33), pp. 3021–3104. doi: 10.1093/eurheartj/ehy339.

Wod, M. *et al.* (2018) ‘Lack of Association Between Proton Pump Inhibitor Use and Cognitive Decline’, *Clinical Gastroenterology and Hepatology*. The American Gastroenterological Association, 16(5), pp. 681–689. doi: 10.1016/j.cgh.2018.01.034.

Wouters, H. J. C. M. *et al.* (2020) ‘Epidemiology of thyroid disorders in the Lifelines Cohort Study (the Netherlands)’, *PLOS ONE*. Edited by S. Naitza, 15(11), p. e0242795. doi: 10.1371/journal.pone.0242795.

Yang, Y. and Metz, D. C. (2010) ‘Safety of Proton Pump Inhibitor Exposure’, *Gastroenterology*. Elsevier Inc., 139(4), pp. 1115–1127. doi: 10.1053/j.gastro.2010.08.023.

5 Acknowledgement

First and foremost, I would like to thank *Monique M.B. Breteler* for imparting the valuable lesson of maintaining a broader perspective, a guiding principle that has been crucial throughout this journey. I am very grateful for her expert guidance and support throughout the course of this dissertation. I would also like to thank my 2nd reviewer *Julia C. Stingl* for her constant support and encouragement as well as her insightful questions and comments. I would also like to thank my 3rd and 4th thesis committee members, *Robin P. Peeters* and *Andreas Mayr*, for their help, cooperation, time and expertise.

I would like to express my heartfelt gratitude to my immediate supervisor *Dianna F.M. de Vries*, for her unwavering support, invaluable guidance, patient mentoring, and persistent efforts over the past four years. This thesis owes its realization to her tireless contributions. *Zonder jouw help zou ik nooit zo ver gekomen zijn. Hiervoor heel erg bedankt!*

I would also like to thank *all the study assistants* and *participants* in the Rhineland Study. Without their contribution and participation, this study would not have been possible.

A big thank you also goes to *all my fellow PhD students* and *colleagues (Dan, Valentina, Santiago, Juliana, Rika, Aslam, Peng, Xingwang, Davide, Weiyi, Konstantinos, Martina, Lucas, Alex, Ahmad)* for fostering such a wonderful working atmosphere. Special thanks to my *Care Bears* for their support and help: *Santiago*, *Rika*, *Dan*, *Valentina*, *Lucas* and *Jack*.

I would like to give special recognition to a few people who have gone from being colleagues to hopefully lifelong friends:

Meta Bönniger, *Fabienne Fox* and *Elvire Landstra*: I am deeply grateful for your friendship, our lunchtime walks together, your invaluable help and support, and for patiently teaching me R.

Annabell Coors, *Ximena Orozco Ruiz*, *Eva Rudholzer* and *Hannah Stocker*: I would like to thank you for your friendship and for being amazing office mates. Your company has brought both joy and camaraderie, with conversations ranging from the light-hearted and humorous to the deep and contemplative.

Gökhan Pehlivan, Valerie Lohner, and *Sophia Tataridis* thank you for your friendship and support. Our coffee breaks together and the many wonderful and humorous moments we shared have made this journey even more memorable.

Thanks also go to the entire *RS-IT team*, especially to *Fabrice Hess* and *Astrid Seidel,* for their help and support.

Of course, I would also like to thank those who have supported and accompanied me over the years: *Benny*, *Darja*, *Philipp*, *Elvin*, *Aylin*, *Gaby*, *Frank*, *Madina*, *Mina*, *Parnaz*, *Victoria*, *Carole*, *Franzi*, *Sebastian, Tabita, Kurt* and *Ute*. You have been a constant source of strength, joy, fun and encouragement.

Finally, I would like to express my deep gratitude to my wife *Imke*, whose unwavering presence over the past 15 years has been a pillar of support, offered with endless love and patience. I am also immensely grateful to my daughter *Mina Juliet*, whose bright spirit has brought light and joy to my days over the past three years. Lastly, I would like to express my heartfelt appreciation and love to my *mother* and *brother* for their enduring love and support, that have been a constant throughout my life. Their presence has been instrumental in shaping my journey.