

Aus der Abteilung SHIP/ Klinisch-Epidemiologische Forschung –  
Institut für Community Medicine  
(Leiter Prof. Dr. med. Henry Völzke)  
der Universitätsmedizin der Universität Greifswald

**Suicidal ideation among participants of population-based studies –  
how to deal with it?**

Inaugural - Dissertation

zur

Erlangung des akademischen Grades

Doktor der Medizin (Dr. med.)

der

Universitätsmedizin

der

Universität Greifswald

2023

vorgelegt von:  
Christiane Günther  
geb. am: 10.04.1994  
in: Anklam

Dekan: Prof. Dr. med. Karlhans Endlich  
1. Gutachter: Prof. Dr. Carsten Oliver Schmidt  
2. Gutachter: Prof. Dr. Steffi G. Riedel-Heller  
Ort, Raum: Universitätsmedizin Greifswald, J 02.16  
Tag der Disputation: 21.03.2024

# Table of Contents

<b>1 Introduction</b> .....	<b>1</b>
1.1 Suicide in research .....	1
1.2 Instruments to assess depressive symptoms and suicidal ideation.....	3
1.3 The relevance of incidental psychiatric findings in observational studies .....	5
1.4 What is missing? .....	6
<b>2 Aims and Objectives</b> .....	<b>7</b>
<b>3 Methods</b> .....	<b>8</b>
3.1 Study design .....	8
3.2 Instruments.....	9
3.2.1 Measurement of depression and depressive symptoms.....	9
3.2.1.1 Munich-Composite International Diagnostic Interview (M-CIDI) .....	10
3.2.1.2 Composite International Diagnostic-Screener (CID-S).....	10
3.2.1.3 Patient Health Questionnaire-9 (PHQ-9).....	10
3.2.1.4 Beck Depression Inventory-II (BDI-II) .....	11
3.2.2 Measurement of suicidal ideation.....	12
3.2.3 Classification of mortality.....	12
3.3 Statistical methods.....	13
3.4 Ethics and data protection .....	13
<b>4 Results</b> .....	<b>14</b>
4.1 Sample baseline characteristics .....	14
4.2 Depression and depressive symptoms .....	14
4.2.1 CID-S .....	14
4.2.2 M-CIDI.....	16
4.2.3 PHQ-9 .....	16
4.2.4 BDI-II .....	16
4.3 Suicidal ideation.....	16
4.4 Mortality and potential suicide .....	18
4.4.1 Potential suicide after indications of depression and depressive symptoms	18
4.4.2 Potential suicide after affirming suicidal ideation.....	20
<b>5 Discussion</b> .....	<b>21</b>
5.1 Main Findings.....	21

5.2 Interpretation .....	21
5.2.1 Depression and depressive symptoms in population-based studies .....	21
5.2.2 Suicidal ideation in population-based studies .....	23
5.2.3 Mortality in population-based studies .....	25
5.2.4 Suicide after suicidal ideation and depressive symptoms in population-based studies .....	26
5.2.5 Should we intervene in case of indications of suicide? .....	28
5.3 Strengths and Limitations .....	29
5.4 Conclusion .....	31
<b>6 Summary .....</b>	<b>32</b>
<b>7 References .....</b>	<b>33</b>
<b>Appendix A .....</b>	<b>45</b>
<b>Appendix B (Instruments) .....</b>	<b>46</b>
<b>List of Abbreviations .....</b>	<b>49</b>
<b>List of Tables .....</b>	<b>50</b>
<b>Eidesstattliche Erklärung .....</b>	<b>51</b>

# 1 Introduction

Feedback of findings to participants of population studies has focused predominantly on results from physical assessments (1-3). Psychological self-reports, routinely collected by interviews or questionnaires, have been largely neglected. Of particular relevance are reports of depression and suicidal ideation, as these can indicate the risk of acute self-harm (4). Using data from the Study of Health in Pomerania (SHIP), we will analyze questionnaires on depressive symptoms, items on suicidal ideation, and mortality data with indications of suicide to develop recommendations for dealing with potentially critical psychological findings within population cohorts.

## 1.1 Suicide in research

Worldwide, every year about 800000 people die of suicide (5). Due to the severity of a suicidal event, predicting self-harm is of enormous importance. It has been shown that self-harming thoughts are associated with an increased risk of subsequent suicidal behavior (6) in psychiatric and non-psychiatric populations (7). In this context identifying such thoughts could be an essential factor.

Another well-studied risk factor for suicidal ideation, suicide attempts, and death is depression and hopelessness, as shown in a meta-analysis of longitudinal studies (8).

In addition, a variety of other risk factors for suicide have been described in longitudinal studies in the general population, such as a past attempted suicide as one of the most decisive risk factors (9, 10), sociodemographic factors (11, 12), behavioral factors (13), physical conditions, such as cardiovascular disease risk factors (14) or cancer (15).

Although extensive research on different risk factors related to suicide has been conducted, the identification of persons with increased suicide risk at the population level is not well understood (16), and the literature to date does not provide clear evidence of the extent to which the process from suicidal ideation to behavior can be reliably be predicted and prevented. Reasons for this include the low base rate of suicide deaths (17). Presumably, as a result, only a few prospective studies have

investigated the occurrence of completed suicide following self-reported depressive symptoms in the general population (15, 25-27).

**Table 1:** Suicidal ideation in population-based studies

<b>Study (Reference)</b>	<b>Instrument</b>	<b>Study design</b>	<b>Sample (sample size n)</b>	<b>Results</b>
Hintikka, 2001 (18)	BDI item	Prospective cohort	Finish population (n = 1593)	Score 1 to 3: Men: 14.7% Women: 9.2%
Palmu, 2020 (19)	BDI-13 item	Prospective cohort	Finish population (n = 4373)	Score 1: 89.5% Score 2: 10.3% Score 3: 0.7% Score 4: 0.1% Score 5: 0.2%
Casey, 2008 (20)	BDI item	Cross-sectional	Five European countries (n = 7950)	Score 0: 90.4% Score 1: 8.6% Score 2: 0.7% Score 3: 0.2%
Ladwig, 2010 (21)	PHQ-9 item	Cross-sectional	Augsburg (Germany) and surrounding districts	Score 1 to 3: 5.3%
Schwerdtfeger, 2020 (22)	PHQ-9 item	Cross-sectional	Different regions across Germany (n = 1980)	Score 0: 94.32% Score 1: 4.72% Score 2: 0.76% Score 3: 0.20%
Ten Have, 2009 (23)	CIDI items	Prospective cohort	Dutch population (n = 4848)	2.7%
Gunnell, 2004 (24)	Single suicidal question	Prospective cohort	British population (n = 2404)	2.3%

BDI, Beck Depression Inventory; PHQ-9, Patient Health Questionnaire 9; CIDI, Composite International Diagnostic Interview.

In prospective studies, it is more common in general population samples to examine self-reported suicide attempts (23) and suicidal ideation (18, 19, 23, 24) as outcome variables instead of a suicide death (Table 1), as they are more likely to be observed, with a lifetime prevalence of 9.2% (suicidal ideation) and 2.7% (suicidal attempt) (12) than suicide death with a global age-standardized suicide rate of 0.01% (5). In addition, risk factors for suicide attempts differ from that for suicide deaths,

particularly concerning sociodemographic factors. While suicide attempts are more frequent at a younger age and among women (28), the suicide death rate increases with age, especially among men (29).

For these reasons, the evidence on the occurrence of suicide following indications of potentially necessary disclosures, like depressive symptoms and suicidal ideation, in population-based research, is sparsely studied to date. In this context, reliable measures of suicide-related symptoms are essential for any population-based screening.

## **1.2 Instruments to assess depressive symptoms and suicidal ideation**

A widely used instrument for assessing a clinical diagnosis of major depression is the Composite International Diagnostic Interview (CIDI) (30). A computerized version of the CIDI, the Munich-Composite International Diagnostic Interview (M-CIDI), has repeatedly been used to assess the prevalence of depressive disorders in population-based studies in Germany (31-35). In addition to standardized interviews, such as the CIDI, screening instruments based on self-reports of depressive symptoms are becoming more common in research and practice. They are time-saving, easy to handle and interpret, and have acceptable validity and reliability (36, 37). In contrast to standardized interviews, such screening tools provide insight into the magnitude of depressive symptoms before the criteria for a clinical diagnosis are met (38).

Two widely used screening questionnaires are the Patient Health Questionnaire-9 (PHQ-9) and the Beck Depression Inventory-II (BDI-II). The responses to the questionnaire items are summed to a total score, with a higher score indicating more severe depression. The PHQ-9 consists of nine items based on the diagnostic criteria of major depressive episodes from the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) (39). It was originally designed for use with primary care patients (39). Since then, it has been used to assess the prevalence of depressive symptoms in population surveys among adults in the United States (40, 41), in a representative population-based adult sample in Germany (38), in a Korean nationally representative adult population (42), and in the general adult population in Poland (43).

In the general population, the PHQ-9 was a suitable screening tool for assessing depressive symptoms (43-47).

The BDI-II is a 21-item self-report measure commonly used to screen for depression (48-51) according to diagnostic criteria listed in the DSM-IV (52).

Screening instruments that have primarily been designed to measure depressive symptoms also include single items that are suitable for assessing suicidal ideation (53). Both the PHQ-9 and the BDI-II also include such items. The PHQ-9 item (item 9) asks about "Thoughts that you would be better off dead or of hurting yourself in some way" over the preceding two weeks (39). So far, its utilization is mainly limited to studies with clinical samples that determine the prevalence of suicidal ideation among outpatients (54-56), patients in the Veterans Health Administration (VHA) system (57), patients with multiple sclerosis (58), HIV (59), cancer (60), arthritis (61) and congestive heart failure (62). Population-based evidence on item 9 in the general population is scarce, and in German regions is limited to cross-sectional studies assessing the prevalence of suicidal ideation (21, 22) (Table 1).

Other researchers used item 9 to determine the suicide risk of study participants. (54-57, 63, 64). There is conflicting evidence about the predictive value of item 9 of the PHQ-9 as a suicide screening tool. On the one hand, several studies in clinical samples demonstrated that severe suicidal ideation, measured by item 9, indicated a higher risk of suicide deaths or suicide attempts (54-57). The American Heart Association (AHA) Science Advisory even recommended evaluating all persons with any kind of suicidal ideation on item 9 (PHQ-9 item 9 score  $\geq 1$  point) for acute suicidality (65).

On the other hand, authors of studies among psychiatric outpatients (63) and patients with a primary diagnosis of a mood disorder (64) criticized the low positive predictive values (5.5% and 28.6%) of item 9. However, due to its high sensitivity (87.6% and 95%), these authors still recommended this item as a suicidal risk screening instrument in clinical settings.

In comparison, a study among primary care patients with elevated depression symptoms reported a lower sensitivity (0.69) of item 9 but still suggested its potential use in primary care due to a specificity of 0.84 (4).



The wording of item 9 of the PHQ-9 might be related to limitations in predictive performance. It has been criticized for measuring not only active thoughts of self-harm but also passive suicidal ideation about whether life is worth living (66). It may also be the reason for a high rate of false-positive rates in clinical samples (67, 68) when participants might agree to passive thoughts without having active suicidal plans. A study among cancer outpatients (60) also indicates this assumption. Only one-third of the respondents who reported suicidal ideation on item 9 was assessed as such in a subsequent clinical interview. Another criticism point is that item 9 does not include past suicide attempts as an essential risk factor for suicide (64).

Regarding the suicide item of the BDI-II, which also assesses suicidal ideation or wishes in the two preceding weeks (52), a study in psychiatric outpatients showed valid results for predicting death by suicide and repeated suicide attempts (69). In the general population, it was shown to be suitable for detecting the prevalence of suicidal ideation in prospective cohort studies in Finland (18, 19) and in a cross-sectional study in five European countries (20) (Table 1).

### **1.3 The relevance of incidental psychiatric findings in observational studies**

Incidental findings in health research are ethical and methodological challenges (70). Scientists need to achieve research aims while at the same time not disregarding the well-being of their study participants. While the handling of incidental physical findings has frequently been addressed due to imaging (1-3, 71-75), the appropriate handling of psychological findings in population-based epidemiologic research has largely been neglected. In contrast to incidental imaging findings, the respondents of depression questionnaires know about their psychological constitution and thus about possible thoughts of self-harm. However, it has been shown that individuals with depression have inadequate coping strategies to deal with such stressful events (76). Therefore, if self-report questionnaires identify the potential risk of suicide, this provides an opportunity and maybe even a necessity for intervention and potential liability.

It is common practice to withhold results from questionnaires or interview assessments from study participants, for example, in the SHIP-study (77). This complies with the researchers' need for non-intervention in observational designs but may pose an ethical problem. Any feedback on potentially critical findings may

alter the natural course of diseases. This may lead to the problem of a biased sample that shows systematic and structural differences in disease progression compared to the general population, thus impairing the generalizability of results (70). However, it would be ethically questionable if subjects with very high depression scores or firmly affirmative answers to questions about suicidal ideation were to commit suicide shortly after taking part in a study after neglecting to intervene for methodological reasons.

#### **1.4 What is missing?**

In suicide research, there is a lack of prospective population-based evidence examining associations between potentially critical indications (high total scores and high scores on the suicide item) of self-harm and short-term subsequent suicide death. Such prospective studies need to focus on shorter periods than has been done in the past (15, 25-27), as the majority of first suicide attempts occurred within one year of the onset of ideation (78).

In addition, because of the low base rate of suicide, such studies should have a large number of participants to obtain generalizable results (17). The commonly used cross-sectional design has the limitation that temporal correlations between outcome and risk factors, in this case, suicide and suicidal ideation, cannot be investigated (79).

There is little evidence of the association between suicidal ideation in the general population and subsequent short-term suicides. Previous research on the use of single suicide items has mainly been limited to studies with clinical or small samples, leading to a lack of generalizability to the general population. This leads to insufficient guidance on recommendations to disclose results from depression instruments in population-based research.

## **2 Aims and Objectives**

In this study, we use data from a prospective population-based cohort study with a large general adult population sample. Our main goal is to provide guidance for dealing with participants of population studies who are classified into a defined depression severity grade, and subjects who strongly affirm single critical suicide ideation items. Therefore, we examine the prediction of short-term suicides based on the provided responses to depression questionnaires and single suicide items. This work focuses on self-harming behavior within one year after study participation, as this might more likely be considered avoidable had critical depression scores been taken into account. In this context, the present study investigates the frequency of depressive symptoms measured by M-CIDI, PHQ-9, BDI-II, Composite International Diagnostic-Screener (CID-S), and the frequency of suicidal ideation measured by suicide items of BDI-II and PHQ-9. By analyzing data from death registries using the International Classification of Diseases, Version 10 (ICD-10) codes, we will also contribute evidence in population-based suicide research.

### **3 Methods**

#### **3.1 Study design**

The Study of Health in Pomerania (SHIP) is a general population cohort study conducted in the region of Western Pomerania in Germany. It targets the description of risk factors (subclinical-) diseases, their consequences, and associations among them. The study includes two independent cohorts. The design of this research project has been described in detail in the study of Völzke et al. (77).

For the baseline SHIP-START-0 assessment, 7008 persons aged 20 to 79 years of German citizenship were selected from a population of 212157 residents from the cities of Greifswald, Stralsund, Anklam, and surrounding 29 communities using a two-stage cluster method. After excluding participants due to death or change of residence, a net sample of 6265 participants remained, of which 4308 subjects (response 68.8%) participated. Further follow-up assessments from 2002 to 2006 (SHIP-START-1), 2007 to 2012 (SHIP-START-2), and 2014 to 2016 (SHIP-START-3) were carried out.

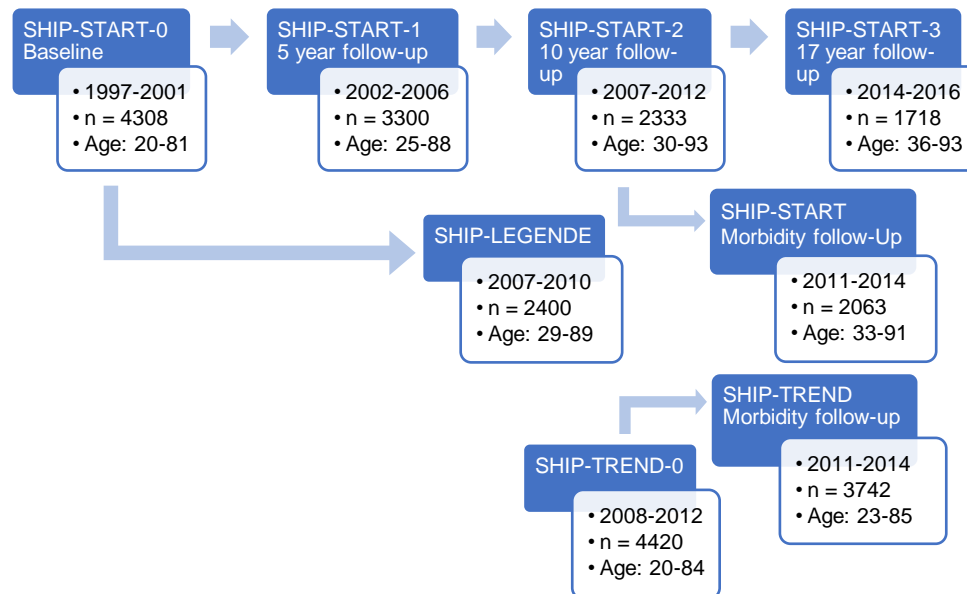
Separately, in parallel to SHIP-START-2, an age- and a sex-stratified randomized sample (SHIP-TREND-0) of 8826 adults with German citizenship (gross sample: 10000) aged 20 to 79 was sampled from the same region. Of these, 4420 (response 50.1%) participated in the baseline survey between 2008 and 2012 (80).

In both samples, contact with study participants was first made by postal mail, then by telephone, and finally, if previously unsuccessful, by a home visit in SHIP-START-0 and SHIP-TREND-0, or via the offer of mobile study centers in SHIP-TREND-0.

SHIP-LEGENDE (Life-Events and Gene-Environment Interaction in Depression) implemented a comprehensive psychological assessment of 2400 participants (gross sample: 3669) from the SHIP-START-0 baseline sample between 2007 and 2010 (81).

Furthermore, a Morbidity follow-up was carried out for SHIP-START and SHIP-TREND from 2011 to 2014. The Morbidity follow-up process included three information sources. Initially, participants were given two postal questionnaires to report any clinical events, which were then validated through a survey of general practitioners. The second source involved obtaining outpatient treatment information from the panel physicians association. Lastly, the University Medicine

Greifswald, the Association of Statutory Health Insurance Physicians, and dentists requested clinical data from SHIP participants.



**Figure 1:** Survey waves of the SHIP-study included in this study (own illustration)

### 3.2 Instruments

#### 3.2.1 Measurement of depression and depressive symptoms

The assessment of depression and depressive symptoms involved the use of both, validated screening questionnaires and a standardized clinical interview, but instruments varied across survey waves (Table 2).

**Table 2:** Overview of instruments used for the measurement of depression and depressive symptoms

Instrument	Cohort
M-CIDI	SHIP-LEGENDE
CID-S	SHIP-START-0
	SHIP-START-1
	SHIP-START-2
	SHIP-START-3
PHQ-9	SHIP-START Morbidity follow-up
	SHIP-TREND-0
	SHIP-TREND Morbidity follow-up
BDI-II	SHIP-LEGENDE

M-CIDI, Munich-Composite International Diagnostic Interview; CID-S, Composite International Diagnostic-Screener; PHQ-9, Patient Health Questionnaire 9; BDI-II, Beck Depression Inventory-II.

### **3.2.1.1 Munich-Composite International Diagnostic Interview (M-CIDI)**

In SHIP-LEGENDE, the comprehensive standardized M-CIDI was used to diagnose unipolar depression (F32x/F33x) within the last four weeks. It is a modified and extended version of the World Health Organization (WHO)-CIDI, which allows a diagnosis according to DSM-IV and ICD-10 criteria (82).

According to various validation studies, the M-CIDI has acceptable to very good psychometric properties (83, 84).

### **3.2.1.2 Composite International Diagnostic-Screener (CID-S)**

Depressive symptoms in SHIP-START-0 and the follow-up waves SHIP-START-1 to SHIP-START-3 were assessed using the CID-S, which is based on the WHO-CIDI for the assessment of DSM-IV and ICD-10 disorders (85). In SHIP-START-0, it was administered as a 12-item self-report questionnaire. Two of these questions assess depressive symptoms with feelings of sadness, loss of interest, tiredness, or a loss of energy almost every day during the lifetime (see Appendix B, Table B.1). In SHIP-START-1, SHIP-START-2, and SHIP-START-3, these questions were part of an interview and covered the period of the past twelve months (see Appendix B, Table B.2).

Each item offers the answer options "Yes" or "No". If both questions were affirmed, this was considered indicative of potential depressive symptomatology.

Based on CID-S questions, a diagnosis of depression cannot be made according to the criteria of the ICD-10 or the DSM-IV. For the two questions on depressive symptomatology, the combined sensitivity is 62.2%, the combined specificity is 77.1%, the combined positive predictive value (PPV) is 19.6%, and the combined negative predictive value (NPV) is 95.9% (85).

For the entire 12-item CID-S questionnaire, the sensitivity is 80.7%, specificity is 55.3%, PPV is 51.6%, and NPV is 85.1% for lifetime disorders (85).

### **3.2.1.3 Patient Health Questionnaire-9 (PHQ-9)**

In SHIP-TREND-0 and the Morbidity follow-up of both cohorts, the PHQ-9 was used. It is based on DSM-IV criteria and comprises nine items that screen for the severity of depressive symptoms in the last two weeks (see Appendix B, Table B.3). Each item has 4-point response options with "Not at all" (score 0), "Several days" (score

1), “More than half the days” (score 2), and “Nearly every day” (score 3). The sum of the scores of the items ranges from 0 to 27 with the following cut-offs (39):

- 1-4: Minimal depression
- 5-9: Mild depression
- 10-14: Moderate depression
- 15-19: Moderately severe depression
- 20-27: Severe depression

A PHQ-9 score of 10 or more had an 88% sensitivity and an 88% specificity to detect depressive symptoms in a general population (39).

Furthermore, a meta-analysis reported a cut-off point of 10 or more for a sensitivity of 80% to 90% to detect major depression (86). The test-retest reliability and internal reliability of the PHQ-9 were excellent (Cronbach’s alpha ranges from 0.86 to 0.89) (39). Evidence from general population studies concludes the reliability and validity of the PHQ-9 as a screening tool for depression (44, 45, 47).

#### **3.2.1.4 Beck Depression Inventory-II (BDI-II)**

The BDI is a self-report scale of depression severity based on DSM-IV criteria (52). The second edition of the BDI (BDI-II), used in SHIP-LEGENDE, includes cognitive, affective, somatic, and vegetative symptoms. The BDI-II contains 21 items, each comprising a 4-point-Likert-type answering scale ranging from 0 (no symptoms) to 3 (an extreme form of each symptom) (see Appendix B, Table B.4).

The scores of the items are then summed up. The assessment of severity of depression is evaluated using the following cut-offs (52):

- 0-13: Minimal range
- 14-19: Mild depression
- 20-28: Moderate depression
- 29-63: Severe depression

The BDI-II assesses symptom severity. However, a high BDI-II score does not justify a diagnosis of depression. Several studies have examined the validity and reliability of BDI-II across different populations and countries, as summarized in the review of Wang and Gorenstein (48). Good internal consistency and test-retest reliability of the BDI-II were consistently shown in the community (49, 50), adolescent and adult clinical outpatients (87), and in adult clinical inpatients (51).

Psychometric testing of the German version showed good internal consistency scores (Cronbach's alpha 0.84) and high concordance with other measures of depressive self-report instruments for different populations (inpatient depressives, unremitted depressives, discharged depressives, community samples, students) (88).

### **3.2.2 Measurement of suicidal ideation**

To assess suicidal ideation, we analyzed item 9 of the PHQ-9 and BDI-II. Both measure the occurrence of suicidal ideation within the last two weeks, and each offers four different response options. Item 9 of the BDI-II asks about suicidal ideation with the following answers: "I don't have any thoughts of killing myself" (score 0), "I have thoughts of killing myself, but would not carry them out" (score 1), "I would like to kill myself" (score 2) and "I would kill myself if I had the chance" (score 3) (52).

The PHQ-9 item asks about "Thoughts that you would be better off dead or of hurting yourself in some way" over the preceding two weeks with the options: "Not at all" (score 0), "Several days" (score 1), "More than half the days" (score 2) and "Nearly every day" (score 3) over the past two weeks (39).

### **3.2.3 Classification of mortality**

We used the ICD-10 codes for mortality causes obtained in mortality follow-ups of the cohorts. We classified all deceased participants in SHIP according to three categories: "Potential Suicide", "Suicide", and "Other causes of death". "Suicide" was only assigned for deceased participants with an ICD-10 code in the range of X60-X84 that is meant to code death caused by suicide explicitly. However, it cannot be ruled out that many suicides may not be perceived as such or were misclassified. Therefore, we assign the class "Potential Suicide" to all deceased participants with ICD-10 codes R96, R98, R99 (death with unknown causes), Y85-Y89 (sequelae from self-harm, accidents or unknown causes), and S00-T88 (Injury, poisoning and certain other consequences of external causes) (89). All other deceased participants were classified as "Other causes of death".



### **3.3 Statistical methods**

Data were analyzed using common descriptive statistics, such as n and percentages for categorical data and mean, standard deviation, minimum, and maximum for continuous data as well as answers on questionnaire rating scales. No inferential statistics were applied, as the goal was not to make inferences about an underlying population but about occurrences of depression-related outcomes and potentially related mortality in the study sample.

### **3.4 Ethics and data protection**

Before the beginning of the study, all participants were informed in detail about SHIP and gave written consent. The participants received no financial compensation beside reimbursement for travel costs.

The study was conducted following the Federal and State Commissioners for Data Protection guidelines and the ethical principles of the Helsinki Declaration. All measurement waves received a positive vote from the Ethics Committee of the University Medicine of Greifswald.

## 4 Results

### 4.1 Sample baseline characteristics

The baseline characteristics of the SHIP-START-0 and SHIP-TREND-0 participants are described in Table 3. Little more than half of the participants were female. The SHIP-TREND-0 cohort was older, with an average age of  $51.96 \pm 15.46$  years, compared to SHIP-START-0 ( $49.80 \pm 16.39$  years). Compared to SHIP-TREND-0 participants, SHIP-START-0 participants were more often widowed (8% vs. 5%) and slightly less likely to be married or in a partnership (75% vs. 78%).

**Table 3:** SHIP-START-0 and SHIP-TREND-0 baseline characteristics

	SHIP-START-0	SHIP-TREND-0
n	4308	4420
Period	1997-2001	2008-2012
Age		
Mean (SD)	$49.80 \pm 16.39$	$51.96 \pm 15.46$
Min	20	20
Max	81	84
Sex		
n (%) female	2192 (50.88)	2275 (51.47)
Marital status		
n (%) single	479 (11)	485 (11)
n (%) married/partnership	3231 (75)	3418 (78)
n (%) apart/divorced	258 (6)	283 (6)
n (%) widowed	323 (8)	221 (5)

n, number; SD, standard deviation.

### 4.2 Depression and depressive symptoms

Table 4 shows the relative frequency of depressive symptoms from the questionnaires (CID-S, PHQ-9, BDI-II) and symptoms of unipolar depression in the last four weeks (F32x/F33x by M-CIDI) of each measurement wave in each of the two cohorts by gender.

#### 4.2.1 CID-S

The proportion of depressive symptoms measured by the CID-S concerning lifetime prevalence in SHIP-START-0 was 12.90%, and concerning 12-month prevalence in

the follow-up waves ranged from 6.70% in SHIP-START-3 to 7.60% in SHIP-START-1.

**Table 4:** Depression (by M-CIDI) and depressive symptoms (by CID-S, PHQ-9 and BDI-II) by gender

		Total		Male		Female	
		n	%	n	%	n	%
<b>M-CIDI</b>							
SHIP-LEGENDE (n = 2383)	F32x/F33x	55	2.31	16	1.41	39	3.12
<b>CID-S</b>							
SHIP-START-0 (n = 4250)	2 symptoms	547	12.90	170	8.10	377	17.40
SHIP-START-1 (n = 3282)	2 symptoms	248	7.60	79	5.00	169	9.90
SHIP-START-2 (n = 2317)	2 symptoms	158	6.80	37	3.40	121	9.90
SHIP-START-3 (n = 1705)	2 symptoms	115	6.70	33	4.20	82	8.90
<b>PHQ-9 category<sup>a</sup></b>							
SHIP-START Morbidity follow-up (n = 1976)	Minimal	1377	69.70	697	76.20	680	64.10
	Mild	444	22.50	160	17.50	284	26.80
	Moderate	104	5.30	35	3.80	69	6.50
	Moderately severe	36	1.80	17	1.90	19	1.80
	Severe	15	0.80	6	0.70	9	0.80
SHIP-TREND-0 (n = 4171)	Minimal	2752	66.00	1459	72.40	1293	59.90
	Mild	1114	26.70	436	21.60	678	31.40
	Moderate	240	5.80	93	4.60	147	6.80
	Moderately severe	45	1.10	20	1.00	25	1.20
	Severe	20	0.50	6	0.30	14	0.60
SHIP-TREND Morbidity follow-up (n = 3545)	Minimal	2383	67.20	1233	73.20	1150	61.80
	Mild	851	24.00	336	19.90	515	27.70
	Moderate	203	5.70	72	4.30	131	7.00
	Moderately severe	74	2.10	30	1.80	44	2.40
	Severe	34	1.00	14	0.80	20	1.10
<b>BDI-II category<sup>b</sup></b>							
SHIP-LEGENDE (n = 2231)	Minimal	1920	86.10	942	89.00	978	83.40
	Mild	190	8.50	77	7.30	113	9.60
	Moderate	79	3.50	29	2.70	50	4.30
	Severe	42	1.90	11	1.00	31	2.60

n, number; M-CIDI, Munich-Composite International Diagnostic Interview; CID-S, Composite International Diagnostic-Screener; PHQ-9, Patient Health Questionnaire 9; BDI-II, Beck Depression Inventory-II

<sup>a</sup> Response categories: Minimal, <5 points; Mild, 5-9 points; Moderate, 10-14 points; Moderately severe, 15-19 points; Severe, 20-27 points

<sup>b</sup> Response categories: Minimal, 0-13 points; Mild, 14-19 points; Moderate, 20-28 points; Severe, 29-63 points.

The proportion of two depressive symptoms (CID-S) in the respective cohort was two to almost three times higher in females than males.

#### **4.2.2 M-CIDI**

In SHIP-LEGENDE, the proportion of subjects fulfilling the criteria for unipolar depression (F32x/F33x) within the last four weeks (M-CIDI) preceding the clinical interview was 2.31%. The proportion was more than twice as high in women (3.12%) compared to men (1.41%).

#### **4.2.3 PHQ-9**

According to the PHQ-9, the frequency of respondents with “Moderate”, “Moderately severe”, and “Severe depression” (PHQ-9 score  $\geq 10$  points) ranged from 7.40% (SHIP-TREND-0) to 8.80% (SHIP-TREND Morbidity follow-up).

Higher severity categories occurred less often, with the largest proportion in the “Minimal depression” category (between 66.00% in SHIP-TREND-0 and 69.70% in SHIP-START Morbidity follow-up) and the lowest proportion in the “Severe depression” category (between 0.50% in SHIP-TREND-0 and 1.00% in SHIP-TREND Morbidity follow-up).

“Minimal depression” occurred more often in men (between 72.40% in SHIP-TREND-0 and 76.20% SHIP-START Morbidity follow-up) compared to women (between 59.90% in SHIP-TREND-0 and 64.10% in SHIP-START Morbidity follow-up). In all other categories (PHQ-9 score  $\geq 5$  points), women were more frequent than men.

#### **4.2.4 BDI-II**

In the BDI-II, employed in SHIP-LEGENDE, the total proportion of “Moderate,” and “Severe depression” (BDI-II score  $\geq 20$  points) was 5.40%. The gender distribution was similar compared with the PHQ-9. On the one hand, men were more often (89.00% vs. 83.40%) in the “Minimal” range category. On the other hand, women were more frequent in each of the other categories (BDI-II  $\geq 14$  points).

### **4.3 Suicidal ideation**

The frequency of suicidal ideation, measured by single items of the PHQ-9 and BDI-II, is shown in Table 5 for each cohort by gender.

According to item 9 of the PHQ-9, the frequency of participants with suicidal ideation (score 1 to 3) in the past two weeks was around 6%. In total, women were more frequently affected by suicidal ideation in all cohorts (women: 6.30% in SHIP-START Morbidity follow-up to 6.90% in SHIP-TREND-0 vs. men: 4.80% in SHIP-START Morbidity follow-up to 5.55% in SHIP-TREND-0). This difference was not evident in the highest categories (scores 2 and 3).

**Table 5:** Suicidal ideation (by single items of PHQ-9 and BDI-II) by gender

	PHQ-9, item 9 score <sup>a</sup>	Total		Male		Female	
		n	%	n	%	n	%
SHIP-START Morbidity follow-up (n = 2020)	0	1898	94.00	880	94.20	1018	93.70
	1	106	5.20	45	4.80	61	5.60
	2	10	0.50	5	0.50	5	0.50
	3	6	0.30	4	0.40	2	0.20
SHIP-TREND-0 (n = 4214)	0	3952	93.80	1922	94.50	2030	93.10
	1	231	5.50	97	4.80	134	6.10
	2	23	0.50	11	0.50	12	0.60
	3	8	0.20	4	0.20	4	0.20
SHIP-TREND Morbidity follow-up (n = 3650)	0	3433	94.10	1647	94.80	1786	93.40
	1	177	4.80	70	4.00	107	5.60
	2	17	0.50	10	0.60	7	0.40
	3	23	0.60	10	0.60	13	0.70
SHIP-LEGENDE (n = 2313)	<b>BDI-II, item 9 score<sup>b</sup></b>						
	0	2159	93.30	1036	94.30	1123	92.50
	1	144	6.20	57	5.20	87	7.20
	2	10	0.40	6	0.50	4	0.30
	3	0	0.00	0	0.00	0	0.00

n, number; PHQ-9, Patient Health Questionnaire 9; BDI-II, Beck Depression Inventory-II

<sup>a</sup> Response categories: 0, Not at all; 1, Several Days; 2, More than half the days; 3, Nearly Every Day

<sup>b</sup> Response categories: 0, I don't have any thoughts of killing myself; 1, I have thoughts of killing myself, but I would not carry them out; 2, I would like to kill myself; 3, I would kill myself if I had the chance.

In SHIP-LEGENDE, 6.60% of respondents to item 9 of the BDI-II reported suicidal ideation within the past two weeks, with the highest score (score 3: "I would kill myself if I had the chance") not being affirmed by any of the participants. In SHIP-LEGENDE, women were more likely to have suicidal ideation than men (7.50% vs.

5.70%), although there was a bigger proportion of men (0.50% vs. 0.30%) scoring more severe (BDI-II, item 9 score 2) expressions of suicidal ideation.

#### 4.4 Mortality and potential suicide

Table 6 shows the mortality and potential suicide of participants in each cohort within and after one year.

Mortality in the SHIP-START-0 cohort (26.62%) was nearly four times as high as in the SHIP-Trend-0 cohort (6.90%).

Within one year after participation, only one case of potential suicide was registered in the SHIP-START-0 cohort. Mortality decreased with each follow-up within the SHIP-START cohort, so it was lowest in SHIP-START-3 (2.91%).

**Table 6:** Mortality and potential suicide by time

	Within one year						After one year			
	Total deaths		Potential suicide		Other causes of death		Potential suicide		Other causes of death	
	n	%	n	%	n	%	n	%	n	%
SHIP-START-0 (n = 4308)	1147	26.62	1	0.02	24	0.56	27	0.63	1095	25.42
SHIP-START-1 (n = 3300)	699	21.18	0	0.00	29	0.88	14	0.42	656	19.88
SHIP-LEGENDE (n = 2400)	296	12.33	0	0.00	12	0.50	10	0.42	274	11.42
SHIP-START-2 (n = 2333)	232	9.94	0	0.00	12	0.51	6	0.26	214	9.17
SHIP-START-3 (n = 1718)	50	2.91	0	0.00	7	0.41	0	0.00	43	2.50
SHIP-START Morbidity follow-up (n = 2063)	145	7.03	0	0.00	16	0.78	4	0.19	125	6.06
SHIP-TREND-0 (n = 4420)	305	6.90	0	0.00	15	0.34	9	0.20	281	6.36
SHIP-TREND Morbidity follow-up (n = 3742)	196	5.24	0	0.00	9	0.51	4	0.11	173	4.62

n, number.

#### 4.4.1 Potential suicide after indications of depression and depressive symptoms

Within one year after the respective SHIP examination, there was no case of suicide or potential suicide in persons with depressive symptoms as measured with the PHQ-9 and BDI-II (PHQ-9 score  $\geq 10$  points, BDI-II score  $\geq 20$  points) or with unipolar depression, as measured with M-CIDI (Table 7).

**Table 7:** Potential suicide after indications of depression (by M-CIDI) and depressive symptoms (by CID-S, PHQ-9, BDI-II) by time

		Within one year				After one year			
		Potential suicide		Other causes of death		Potential suicide		Other causes of death	
		n	%	n	%	n	%	n	%
	<b>M-CIDI (F32x/F33x)</b>								
SHIP-LEGENDE	n = 55	0	0.00	1	1.82	0	0.00	2	3.64
	<b>CID-S (2 symptoms)</b>								
SHIP-START-0	n = 547	1	0.18	1	0.18	3	0.54	103	18.83
SHIP-START-1	n = 248	0	0.00	6	2.42	5	2.02	43	17.34
SHIP-START-2	n = 158	0	0.00	3	1.90	2	1.30	14	8.90
SHIP-START-3	n = 115	0	0.00	0	0.00	0	0.00	4	3.50
	<b>PHQ-9 category (≥ 10 points)<sup>a</sup></b>								
SHIP-START Morbidity follow-up	Moderate (n = 104)	0	0.00	1	1.00	1	0.96	10	9.60
	Moderately severe (n = 36)	0	0.00	1	2.80	0	0.00	4	11.10
	Severe (n = 15)	0	0.00	0	0.00	0	0.00	3	20.00
SHIP-TREND-0	Moderate (n = 240)	0	0.00	2	0.83	2	0.83	13	5.40
	Moderately severe (n = 45)	0	0.00	1	2.20	0	0.00	6	13.30
	Severe (n = 20)	0	0.00	1	5.00	0	0.00	1	5.00
SHIP-TREND Morbidity follow-up	Moderate (n = 203)	0	0.00	2	1.00	1	0.49	15	7.40
	Moderately severe (n = 74)	0	0.00	0	0.00	0	0.00	4	5.40
	Severe (n = 34)	0	0.00	1	2.90	0	0.00	4	11.80
	<b>BDI-II category (≥ 20 points)<sup>b</sup></b>								
SHIP-LEGENDE	Moderate (n = 79)	0	0.00	1	1.30	2	2.50	7	8.90
	Severe (n = 42)	0	0.00	0	0.00	1	2.40	8	19.00

n, number; M-CIDI, Munich-Composite International Diagnostic Interview; CID-S, Composite International Diagnostic-Screener; PHQ-9, Patient Health Questionnaire 9; BDI-II, Beck Depression Inventory-II

<sup>a</sup> Response categories: Moderate, 10-14 points; Moderately severe, 15-19 points; Severe, 20-27 points

<sup>b</sup> Response categories: Moderate, 20-28 points; Severe, 29-63 points.

There was one case in SHIP-START-0 of potential suicide within one year after reporting two depressive symptoms in the CID-S.

Within the category "Other causes of death within one year", the proportion of those with "Moderate depression" by PHQ-9 and BDI-II was larger than in the category of "Severe depression".

Suicide more than one year after answering the PHQ-9 and BDI-II could not be excluded in six cases with "Moderate depression" and one case (in SHIP-LEGENDE) with "Severe depression".

#### 4.4.2 Potential suicide after affirming suicidal ideation

There was no case of a participant with suicidal ideation according to item 9, PHQ-9 or item 9, BDI-II who died of suicide or potential suicide within one year after responding (Table 8). Therefore this category is not present in Table 8. In five cases, potential suicide occurred more than one year after the SHIP examination. Three of these participants reported suicidal ideation on "Several Days" and two on "More than half the days" on item 9 of the PHQ-9.

**Table 8:** Potential suicide and other causes of death after indications of suicidal ideation (by single items of PHQ-9 and BDI-II) by time

	PHQ-9, item 9 score <sup>a</sup>	Within one year		After one year			
		Other causes of death		Potential suicide		Other causes of death	
		n	%	n	%	n	%
SHIP-START Morbidity follow-up	1 (n = 106)	3	2.80	0	0	10	9.40
	2 (n = 10)	0	0	1	10.00	3	30.00
	3 (n = 6)	0	0	0	0	1	16.70
SHIP-TREND-0	1 (n = 231)	2	0.90	2	0.90	11	4.80
	2 (n = 23)	0	0	0	0	1	4.30
	3 (n = 8)	0	0	0	0	1	12.50
SHIP-TREND Morbidity follow-up	1 (n = 177)	0	0	1	0.60	12	6.8
	2 (n = 17)	1	5.90	1	5.88	2	11.80
	3 (n = 23)	0	0	0	0	2	8.70
SHIP-LEGENDE	<b>BDI-II, item 9 score<sup>b</sup></b>						
	1 (n = 144)	0	0	0	0	20	13.90
	2 (n = 10)	0	0	0	0	1	10.00
	3 (n = 0)	0	0	0	0	0	0

n, number; PHQ-9, Patient Health Questionnaire 9; BDI-II, Beck Depression Inventory-II

<sup>a</sup> Response categories: 1, Several Days; 2, More than half the days; 3, Nearly Every Day

<sup>b</sup> Response categories: 1, I have thoughts of killing myself, but I would not carry them out; 2, I would like to kill myself; 3, I would kill myself if I had the chance.



## **5 Discussion**

### **5.1 Main Findings**

This work aims to provide guidance on whether severe self-reported depressive findings should be communicated after participation in population-based cohort studies to prevent short-term suicides. Therefore, premature deaths within one year after SHIP examination were assessed in participants who provided potentially critical indications from a clinical interview, depression questionnaires, and single items on suicidal ideation.

In conclusion, we had no evidence of suicidal events, regardless of depression scores or responses to specific items related to suicidal ideation, within one year after participation in our population-based sample. This was the case despite the reasonably high proportion of depressive symptoms among the participants, with women being affected more frequently than men overall, especially in the higher severity categories.

Therefore, from our results, we cannot conclude that even after self-reported severe depression-related symptoms from questionnaires or clinical interviews, findings should be reported back to participants or further diagnostics be triggered to prevent suicides in a population-based observational setting proactively.

### **5.2 Interpretation**

#### **5.2.1 Depression and depressive symptoms in population-based studies**

When interpreting our results, one central aspect of the use of findings beyond SHIP is whether results on depression and depressive symptoms are in line with findings from other cohorts. Within the German Health Interview and Examination Survey (DEGS1) (38), which was conducted among adults aged 18-79, the frequency of current depressive symptoms (PHQ-9  $\geq$  10 points) was 8.1% (38) and thus similar to our results. As in our study, the frequency of affected women was higher than for men (10.2% vs. 6.1%) for recommended cut-off scores (PHQ-9  $\geq$  10 points).

The US National Health and Nutrition Examination Survey (NHANES) (40) also reported a similar frequency of 6.8% (NHANES) using the same cut-off of  $\geq$  10 points of the PHQ-9. A nationwide, population-based survey of the Korean adult general population also showed comparable distributions of the frequency of 4.4% for “Moderate”, 1.5% for “Moderately severe”, and 0.6% for “Severe depression”

(42). In a study of the general population in Poland (43), the sum of the frequency for “Moderate”, “Moderate to severe”, or “Severe depression” was 18%, which is more than twice as high as that in our study, with women being also more frequently prevalent in these categories than men (22.2% vs. 12.2%). This may be due to different population- and sample characteristics with a significantly smaller age range (35 to 64 years) and a larger proportion of women in the Polish study (43). In Germany, only one study uses the BDI-II as a measurement tool for depressive symptoms in a non-clinical sample (88). The mean BDI-II total score of 7.69 (88) was comparable to 6.41 in SHIP-LEGENDE.

Overall, findings from other studies indicate that our results are in a plausible range. The 12-month prevalence of depression in other population-based studies in Germany (31, 32), measured by the M-CIDI, was 7.7% and 10.7%. Due to the different study periods, these results are higher than the 1-month prevalence (2.31%) in SHIP-LEGENDE in our study. Higher prevalences with more extended study periods also show results from studies within SHIP-LEGENDE that examined the 12-month prevalence of major depressive disorder (5.0%) (35) and the lifetime prevalence of unipolar depression (women: 22.8%, men: 11.3%) (33). Similar to our results, in these studies, women were affected considerably more frequently by depression.

The approximately half as high proportions of depressive symptoms in SHIP-START-1, SHIP-START-2 and SHIP-START-3 compared to the baseline cohort can be explained by the fact that in these follow-up waves, the CID-S assessed the period over the last twelve months. In contrast, the CID-S in SHIP-START-0 measured the lifetime prevalence.

In general, comparisons between the different instruments are of limited value, as it has been shown in a large population-based study in Germany that prevalence rates vary by the type of assessment instrument and diagnostic thresholds (90). In this study, the prevalence of depression measured by PHQ-9 was also lower than that measured by CIDI, and the agreement between both instruments was moderate (Cohen's kappa: 0.43). This is plausible, considering that PHQ-9 covers the narrow time frame of the last two weeks.

Within our study, the distribution of the BDI-II and PHQ-9 scores in the different survey waves shows similar tendencies. The largest proportion of participants was

in the "Minimal" category (from 66.00% in SHIP-Trend-0 to 86.10% in SHIP-LEGENDE), and the smallest proportion in the category of "Severe depression" (0.50% in SHIP-Trend-0 to 1.90% in SHIP-LEGENDE). This is a decisive difference from studies with clinical samples, in which patients were significantly more likely to be affected by "Severe depression" based on PHQ-9, with a frequency of 17.2% (57) and 13% (54), respectively. A possible explanation may be that individuals with severe depressive symptoms are probably less likely to participate in population-based health surveys. It has been shown that people with psychiatric and chronic diagnoses in population-based studies participated less in studies than people without such diagnoses (91).

Moreover, depression is associated with morbidity (92). Therefore, we assume that particularly ill and depressed participants no longer participate in follow-up studies. This results in subsamples with fewer depressive symptoms among the participants compared to the baseline. This suggests that people with severe depressive symptoms may be underrepresented in the follow-up cohorts.

### **5.2.2 Suicidal ideation in population-based studies**

SHIP is one of only a few population-based prospective cohort studies (18, 19, 23, 24), see Table 1 (page 2), to report suicidal ideation. In two of these (18, 19), the authors used an item of a BDI version (BDI-21 or BDI-13) comparable to our study in Finnish general population samples. In the study of Hintikka et al. (18), the prevalence of all types of responses for suicidal ideation (score 1 to 3), excluding the response "I don't have any thoughts of harming myself" (score 0), was 11.5%. The frequency is thus nearly twice as high as in SHIP-LEGENDE (6.60%). However, compared with SHIP-LEGENDE, the Finnish cohort was younger on average (25 to 64 years), with the highest proportion of suicidal ideation occurring within the youngest age group (25 to 34 years). These results are not surprising because younger age is a risk factor for suicidal ideation (28). Palmu et al. (19) explored suicidal ideation in a sample of comparable mean age (54.8 years) to SHIP-LEGENDE (55.9 years). An item of the BDI-13 with five alternative statements of suicidal ideation was used, resulting in similar patterns to our study. Thus, the frequency was 89.5% for the lowest category: "I have never thought to harm myself", and 0.2% for the highest category: "I wish I would be dead".

In two of the other population-based prospective cohort studies (23, 24), suicidal ideation was assessed by different instruments, which limits the comparability to our results. In the first one, a 3-wave cohort study in the Dutch adult general population (23) of a comparable sample size ( $n = 4848$ ) to SHIP-TREND-0 ( $n = 4214$ ), the researcher used CIDI items on suicidality. The 3-year incidence of suicidal ideation was 2.7%. This is lower than the 2-week prevalence of suicidal ideation in SHIP-TREND-0 (6.20%). In the second one, an 18-month follow-up cohort study in the British population (24) with a comparable sample size ( $n = 2404$ ) to SHIP-LEGENDE ( $n = 2313$ ) with an age range of 16 to 74 years, the authors reported an annual incidence of suicidal ideation of 2.3%. They measured suicidal ideation by the single question: "Have you thought of taking your life, even if you would not really do it?" (24), which is comparable to BDI-II, item 9 score 1: "I have thoughts of killing myself, but I would not carry them out". This result is lower than the one for BDI-II, item 9 score 1 in SHIP-LEGENDE (6.20%). We explain these differences by the use of different assessment tools and different metrics.

In another cross-sectional study (20), the authors used item 9 of the BDI in general population samples from five European countries ( $n = 7950$ ) aged 18 to 64. They reported a similar distribution of the various severities of suicidal ideation compared to our cohorts. Severe suicidal ideation, measured by the responses "I would like to kill myself" (score 2) and "I would kill myself if I had the chance" (score 3), occurred with a mean prevalence of 0.9% considerably less frequently than the mild form of suicidal ideation: "I have thoughts of killing myself, but I would not carry this out" (score 1) with 8.6% (20). As in our study, the participants in the British and urban Irish samples did not choose the most severe expression on item 9 of the BDI (score 3). In these two samples, the gender distribution was comparable to our results, with a bigger proportion of women scoring milder (BDI-II, item 9 score 1) and a bigger proportion of men scoring more severe (BDI-II, item 9 score 2) expressions of suicidal ideation. As in SHIP-START Morbidity follow-up and SHIP-LEGENDE, men were even more frequent in this range than women. This could be because men seem to be affected by more severe suicide ideation, as it has been shown that they use methods for concrete suicide attempts and suicides more often than women (11).

Concerning item 9 of the PHQ-9, our results show comparable frequencies of suicidal ideation to studies in the general population in Germany (21, 22). Ladwig et al. (21) reported a prevalence of suicidal ideation of any severity (score 1 to 3) of 5.3% in their population-based study from the city of Augsburg and surrounding districts with 3079 subjects aged 35 to 84 years compared to 6.20% in SHIP-Trend-0. The study of Schwerdtfeger et al. within different regions across Germany (22) consisted of a smaller (n = 1980) and younger (mean age: 46.6 years) sample than SHIP-TREND-0. Nevertheless, the frequency of suicidal ideation, with a total of 5.7% for any severity of suicidal ideation, was comparable to this in SHIP-TREND-0 (6.20%).

### **5.2.3 Mortality in population-based studies**

Our results show that with each follow-up wave, mortality decreases significantly and is about ten times lower in SHIP-START-3 (2.91%) compared to the baseline cohort (26.62%). This is because the data were collected from the beginning of each wave until the end of the observation period. With each follow-up wave, the follow-up time is shorter. This also explains the approximately fourfold higher mortality (26.62%) in SHIP-START-0, compared to SHIP-Trend-0 (6.90%), in which the follow-up time was considerably shorter.

Considering the mortality within the first year after participation in the study, the mortality was very low compared to the age-specific mortality per year within the population of Mecklenburg-Western Pomerania, which would suggest considerably higher death rates (see Appendix A, Table A.1). For example, for SHIP-START-2, based on the distribution in a similar age range as in SHIP-START-2, we would expect a mortality of 20 per 1000 participants per year (93), but in the first year after participation, there were only 12 deaths among 2333 participants in SHIP-START-2. Similar observations were reported in other population-based prospective follow-up studies, which have shown lower mortality rates among cohort study participants than non-participants (91, 94-96). This applied even to follow-up examinations 15 years (96) or 20 years (95) after the baseline examination. It has been suggested (91) that particularly people with severe chronic or psychiatric illnesses were less likely to participate in population-based studies. Presumably, after each SHIP wave, this "filter" of the particularly ill participants occurred again, resulting in a robust

subsample concerning health status. This could explain the low mortality in the year after participation in SHIP.

The low base rate of suicides and potential suicides in our study is plausible, with a total of one case among all cohorts in the total observation period, considering the low suicide rate in Mecklenburg-Western Pomerania in 2020 (0.01%) (97).

#### **5.2.4 Suicide after suicidal ideation and depressive symptoms in population-based studies**

This is, to the best of our knowledge, the first study to examine the occurrence of suicide within one year, after reporting suicidal ideation and after measuring depressive symptoms in a prospective population-based cohort study. We could identify only one potential suicide-related death in this period, even among those with severe suicidal ideation or depressive symptoms.

Deaths following self-reported depressive symptoms have been reported in population-based cohort studies in the general population (15, 25-27), yet based on longer follow-up periods. Furthermore, none of these studies used any of the instruments used in our study to assess depressive symptoms, which leads to limited comparability with our results.

Nevertheless, the findings of these studies show a similar trend to our study: Suicidal deaths were seldom in all of these studies, even among subjects with depression (or depressive disorders) or suicidal ideation. For example, in a cohort study among 3056 participants from the Netherlands (26) only one of 124 persons with a diagnosed depression (measured by “National Institute of Mental Health Diagnostic Interview Schedule” according to DSM-III criteria) committed suicide during the 4-year-follow-up. In another study among 57897 adults aged 25 years and older in the United States with a 2.5-year follow-up (25), two suicide deaths occurred among 615 “depressed” (measured by one single question) participants. In a cohort study among Swedish adults (15), 0.6% of women and 1.4% of men among all 211212 persons with depression (classified according to ICD-10 codes) died of suicide during the observation period of eight years. In a general population study in South Korea (27), 57 completed suicides occurred among 5152 subjects with depressive disorders (classified according to ICD-10 codes) within twelve years.

Inaccurate registration of deaths based on the ICD-10 criteria may contribute to a systematic underrating of suicide deaths (98). Some suicides are misclassified, such as poisoning accidents, which has already led to an underestimation of suicide rates between 10% and 30% (99-101). To counter this possibility, we also analyzed ICD-10 codes suggestive of potential suicides in addition to ICD-10 codes "Intentional self-harm".

In contrast to our findings, studies from clinical settings showed that higher scores on item 9 were associated with significantly elevated suicide risk within one year (57) and 30 days (54) after assessment. However, concerning the large proportion of patients with severe suicidal ideation as measured by "More than half the days" and "Nearly every day" (item 9 score  $\geq 2$ ), suicide rates were relatively low. Thus, the study among patients in the VHA system (57) showed that 31 (0.18%) among 17287 patients with item 9 score  $\geq 2$  died from suicide deaths within one year. The study in outpatients in a large integrated health system in Washington and Idaho (54) revealed six suicide deaths (0.05%) among 11459 patients with item 9 score  $\geq 2$  within 30 days, according to the most recent PHQ questionnaire.

In addition to item 9, the authors of the studies mentioned above also applied the PHQ-9 total score for assessing depressive symptoms. Similarly, the rates of suicide among patients with "Moderate" to "Severe depression" (PHQ-9 score  $\geq 10$ ) were relatively low, being 0.13% (57) and 0.05% (54), respectively. Because of the clinical setting of these studies, these results are of limited comparability with those of our population-based study. Besides, the samples differed in size and gender- and age distribution compared to our cohorts. However, most of those who reported suicidal ideation or depressive symptoms did not commit suicide. This is consistent with our results showing that even those participants who reported suicidal ideation did not commit suicide the following year after participation.

A suspected reason for these findings may be the wording of item 9 (68). The item consists of two parts: the passive thoughts of death and the active thoughts of self-harm. In a medical setting among cancer outpatients (60) and even among patients with depression (102), it was found that patients who responded positively to item 9 seemed to agree more with the passive part. Our results suggest that this assumption may have also played a role in our study since none of the participants committed suicide in the following year after a positive response on item 9. Thus,

we find no need to classify participants of population-based prospective cohort studies as at elevated risk to commit suicide only based on a positive response to individual suicide items.

### **5.2.5 Should we intervene in case of indications of suicide?**

From an ethical aspect, explicit indications of short-term self-harm, recognized as requiring urgent treatment, must be communicated. Our results show no evidence to recommend routine intervention or reporting of findings for such indications in population-based cohort studies. However, we cannot rule out the possibility that there may be constellations of several covariates that are more robust indications of the information provided by the persons for whom communication is necessary. If these indications could be identified, there would be an obligation to report findings. However, we give our recommendations based on the mission of an observational study, not because we deny the importance of preventing even a single potential suicide, but because our study does not have a service mission where it would be necessary to communicate or intervene when participants indicate evidence of severe depression or suicidal ideation. With such intervention, the SHIP-study as an observational study would endanger its purpose and thus lose its legitimacy.

In our study, there was one case of potential suicide (SHIP-START-0) within one year of participating. This person was conspicuous in the screening using CID-S, but this could include many false-positive participants as it asks about depressive symptoms ever in life. Based on PHQ-9 and BDI-II, which measure periods of the past two weeks and include explicit items about suicidal ideation, there was no case of suicide and potential suicide within one year of participation in the entire study. Therefore, we suspect that the one case of potential suicide was not related to depression. In addition, we cannot definitively be sure this case was indeed a potential suicide because of a potential misclassification by death certificate data. Based on our findings, we suggest that individuals with suicidal ideation and severe depressive symptoms who participate in a population-based cohort study do not commit suicide within one year. This suggests that focusing on scores from depression questionnaires or single suicide items alone may lead to misclassifying of participants as suicidal. Therefore, suicide prevention interventions based only



on such scores would likely be ineffective. Concerning suicide risk within one year, based on our results, both suicidal ideation and severe depressive symptomatology or depression can be considered incidental findings when screening with depression questionnaires in prospective cohort studies.

Prevention of long-term risk is not the responsibility of an observational study. The cases of potential suicide more than one year after data collection largely had only mild depressive symptoms or low scores in the suicide item. This suggests that predicting of long-term suicide risk (after one year) using only single screening methods is not reasonable.

### **5.3 Strengths and Limitations**

The major strength of the study is the comprehensive assessment of depression, depressive symptoms, suicidal ideation, and suicides in a large population-based cohort study with long follow-ups. The M-CIDI allows a standardized diagnosis of unipolar depression. In contrast, the use of screening instruments for depressive symptoms (CID-S, BDI-II, PHQ-9) in our study may have led to uncertainties regarding the correct diagnosis, despite good validity (85). The inconsistent use of instruments across survey waves also limits comparisons of results within the cohorts. The demographic and socioeconomic characteristics of the SHIP-study populations roughly reflect those of the average population of Central Europe in general.

However, a limitation of the study is that only residents from the cities of Greifswald, Stralsund, Anklam, and surrounding 29 communities were surveyed. Therefore, generalizability of the results outside this region is limited. However, this work differs significantly from others because it provides evidence about study participants rather than an underlying population. We aim to give recommendations for dealing with critical indications of study participants, and therefore the generalizability of the results is not the focus of this work.

Selection bias due to the selective participation of healthier individuals than the general population cannot be excluded. It may have led to an underrepresentation of depressed and suicidal participants in our sample.

Because BDI-II and PHQ-9 are based on self-report, due to social desirability bias, underestimated prevalence rates of depressive symptomatology and suicidal ideation may have occurred (103).

Another limitation is that suicidality was measured only using single items of the BDI-II and PHQ-9. However, these items do not ask about specific suicide plans or past suicide attempts, which are indicators of later suicidal behavior in the general population (104).

In addition, recall bias when reporting possible symptoms within the past two weeks may have led to underestimating the frequencies of suicidal ideation and depressive symptoms.

Another reason for underestimating frequencies could be that participants with episodic and short-term courses of depression or suicidal ideation were asymptomatic at the time of the survey.

Although the response rate of our study is acceptable, we cannot exclude a possible influence of non-response bias and therefore suspect that the number of depressive and suicidal subjects might be higher among non-participants than among participants.

In addition, we must consider that we may have underestimated the frequencies of suicides, as they may not have been registered as such due to possible misclassification by death certificate data. Despite our large sample size, there was only one case of potential suicide within a year, leading to a limited assessment of an association between suicidal ideation and suicide. However, this is a general problem in research on suicide death in prospective cohort studies. It is challenging to study because this is a low base rate event (105, 106). This is also evident when looking at the annual suicide rate in Mecklenburg-Western Pomerania, with 0.01% in 2020 (97). The low base rate of suicide means that the proportion of suicide deaths within a cohort will always be low, even among participants having potential risk factors, such as suicidal ideation or depressive symptoms. Therefore, it is plausible that we did not observe a suicide in the observation period.

## **5.4 Conclusion**

Within one year of participation in the study, we found no evidence of suicides among SHIP participants. This was the case despite a relatively high proportion of depressive symptoms, with women being more affected than men, especially in the more severe categories.

Therefore, based on our findings, in the context of a general population study, we cannot recommend sharing the results of self-reported suicidal ideation or depressive symptoms to prevent a short-term risk of self-harm, as this needs to be weighed against potential intervention effects due to communicated findings.

Because of our limited case number, we cannot entirely rule out the possibility of self-harm under certain conditions. A repetition of our approach in a larger study would therefore be recommendable.

Furthermore, we did not study the long-term course and did not state that the self-reported symptoms are of no relevance for the studies individuals. However, it must be emphasized that preventing individual long-term health risks among its participants is not the primary responsibility of a longitudinal observational study.

## 6 Summary

Our study examined whether potentially critical indications from depression questionnaires, interviews, and single items on suicidal ideation among participants in a large prospective population-based study are related to short-term suicides within one year. For this purpose, we studied the association between (a) the severity of depressive symptoms according to the M-CIDI and the PHQ-9, BDI-II, and CID-S depression screening and (b) elevated scores on single suicidal ideation items and mortality according to claims databases.

In the baseline cohort, the frequency of depressive symptoms measured by CID-S was 12.90% (SHIP-START-0). The frequency for “Moderate” to “Severe depression” measured by the PHQ-9 ( $\geq 10$  points) and BDI-II ( $\geq 20$  points) questionnaires ranged from 5.40% (SHIP-LEGENDE) to 8.80% (SHIP-TREND Morbidity follow-up). The 1-month prevalence of unipolar depression, measured by the M-CIDI in SHIP LEGENDE, was 2.31%.

Between 5.90% (SHIP-TREND Morbidity follow-up) and 6.60% (SHIP-LEGENDE) of respondents showed a certain degree of suicidal ideation in the two weeks preceding the assessment, according to BDI-II and PHQ-9.

Our results show the high frequency of depressive symptoms in the study region, with women being affected more frequently than men, especially in the higher categories. Furthermore, women were more frequently affected by suicidal ideation, although this difference was not evident in the highest categories.

There was one potential suicide in the year after a SHIP examination.

From our results, we cannot conclude that severe self-reported symptoms from depression questionnaires should be reported back to participants of an observational population-based study to prevent suicide deaths within one year.

## 7 References

1. Hegenscheid K, Seipel R, Schmidt CO, Volzke H, Kuhn JP, Biffar R, et al. Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. *Eur Radiol.* 2013;23(3):816-26.
2. Schmidt CO, Hegenscheid K, Erdmann P, Kohlmann T, Langanke M, Volzke H, et al. Psychosocial consequences and severity of disclosed incidental findings from whole-body MRI in a general population study. *Eur Radiol.* 2013;23(5):1343-51.
3. Schmidt CO, Sierocinski E, Hegenscheid K, Baumeister SE, Grabe HJ, Volzke H. Impact of whole-body MRI in a general population study. *Eur J Epidemiol.* 2016;31(1):31-9.
4. Uebelacker LA, German NM, Gaudiano BA, Miller IW. Patient health questionnaire depression scale as a suicide screening instrument in depressed primary care patients: a cross-sectional study. *The Primary Care Companion For Cns Disorders.* 2011;13(1):1-6.
5. WHO. Suicide in the world: global health estimates. Geneva, Switzerland: World Health Organization; 2019. 7-16 p.
6. Ribeiro JD, Franklin JC, Fox KR, Bentley KH, Kleiman EM, Chang BP, et al. Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. *Psychol Med.* 2016;46(2):225-36.
7. Hubers AAM, Moaddine S, Peersmann SHM, Stijnen T, van Duijn E, van der Mast RC, et al. Suicidal ideation and subsequent completed suicide in both psychiatric and non-psychiatric populations: a meta-analysis. *Epidemiology and psychiatric sciences.* 2018;27(2):186-98.
8. Ribeiro JD, Huang X, Fox KR, Franklin JC. Depression and hopelessness as risk factors for suicide ideation, attempts and death: meta-analysis of longitudinal studies. *Br J Psychiatry.* 2018;212(5):279-86.
9. Suominen K, Isometsä E, Suokas J, Haukka J, Achte K, Lönnqvist J. Completed suicide after a suicide attempt: a 37-year follow-up study. *Am J Psychiatry.* 2004;161(3):562-3.

10. Wang AG, Mortensen G. Core features of repeated suicidal behaviour: a long-term follow-up after suicide attempts in a low-suicide-incidence population. *Soc Psychiatry Psychiatr Epidemiol.* 2006;41(2):103-7.
11. Mejías-Martín Y, Luna Del Castillo JD, Rodríguez-Mejías C, Martí-García C, Valencia-Quintero JP, García-Caro MP. Factors Associated with Suicide Attempts and Suicides in the General Population of Andalusia (Spain). *Int J Environ Res Public Health.* 2019;16(22):4496.
12. Nock MK, Borges G, Bromet EJ, Alonso J, Angermeyer M, Beautrais A, et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br J Psychiatry.* 2008;192(2):98-105.
13. Schneider B, Lukaschek K, Baumert J, Meisinger C, Erazo N, Ladwig KH. Living alone, obesity, and smoking increase risk for suicide independently of depressive mood findings from the population-based MONICA/KORA Augsburg cohort study. *J Affect Disord.* 2014;152-154:416-21.
14. Jee SH, Kivimaki M, Kang HC, Park IS, Samet JM, Batty GD. Cardiovascular disease risk factors in relation to suicide mortality in Asia: prospective cohort study of over one million Korean men and women. *Eur Heart J.* 2011;32(22):2773-80.
15. Crump C, Sundquist K, Sundquist J, Winkleby MA. Sociodemographic, psychiatric and somatic risk factors for suicide: a Swedish national cohort study. *Psychol Med.* 2014;44(2):279-89.
16. Paris J. Predicting and preventing suicide: do we know enough to do either? *Harv Rev Psychiatry.* 2006;14(5):233-40.
17. O'Connor RC, Portzky G. Looking to the Future: A Synthesis of New Developments and Challenges in Suicide Research and Prevention. *Front Psychol.* 2018;9:2139.
18. Hintikka J, Pesonen T, Saarinen P, Tanskanen A, Lehtonen J, Viinamäki H. Suicidal ideation in the Finnish general population. A 12-month follow-up study. *Soc Psychiatry Psychiatr Epidemiol.* 2001;36(12):590-4.
19. Palmu R, Koskinen S, Partonen T. Suicidality in relation to depressive symptoms and psychological distress in adults aged 30 to 101 years in a population-based study in Finland. *Psychiatry Res.* 2020;284:112704.

20. Casey P, Dunn G, Kelly BD, Lehtinen V, Dalgard OS, Dowrick C, et al. The prevalence of suicidal ideation in the general population: results from the Outcome of Depression International Network (ODIN) study. *Soc Psychiatry Psychiatr Epidemiol.* 2008;43(4):299-304.
21. Ladwig KH, Klupsch D, Meisinger C, Baumert J, Erazo N, Schneider A, et al. Gender differences in risk assessment of death wishes and suicidal ideation in the community: results from the KORA Augsburg F3 study with 3079 men and women, 35 to 84 years of age. *J Nerv Ment Dis.* 2010;198(1):52-8.
22. Schwerdtfeger KA, Bahramsoltani M, Spangenberg L, Hallensleben N, Glaesmer H. Depression, suicidal ideation and suicide risk in German veterinarians compared with the general German population. *Vet Rec.* 2020;186(15):e2.
23. Ten Have M, De Graaf R, van Dorsselaer S, Verdurmen J, Van 't Land H, Vollebergh W, et al. Incidence and course of suicidal ideation and suicide attempts in the general population. *Can J Psychiatry.* 2009;54(12):824-33.
24. Gunnell D, Harbord R, Singleton N, Jenkins R, Lewis G. Factors influencing the development and amelioration of suicidal thoughts in the general population. Cohort study. *Br J Psychiatry.* 2004;185:385-93.
25. Zheng D, Macera CA, Croft JB, Giles WH, Davis D, Scott WK. Major depression and all-cause mortality among white adults in the United States. *Ann Epidemiol.* 1997;7(3):213-8.
26. Penninx BW, Geerlings SW, Deeg DJ, van Eijk JT, van Tilburg W, Beekman AT. Minor and major depression and the risk of death in older persons. *Arch Gen Psychiatry.* 1999;56(10):889-95.
27. Lee H, Myung W, Lee C, Choi J, Kim H, Carroll BJ, et al. Clinical epidemiology of long-term suicide risk in a nationwide population-based cohort study in South Korea. *J Psychiatr Res.* 2018;100:47-55.
28. Borges G, Nock MK, Haro Abad JM, Hwang I, Sampson NA, Alonso J, et al. Twelve-month prevalence of and risk factors for suicide attempts in the World Health Organization World Mental Health Surveys. *J Clin Psychiatry.* 2010;71(12):1617-28.
29. Shah A, Bhat R, Zarate-Escudero S, DeLeo D, Erlangsen A. Suicide rates in five-year age-bands after the age of 60 years: the international landscape. *Aging Ment Health.* 2016;20(2):131-8.

30. Haro JM, Arbabzadeh-Bouchez S, Brugha TS, de Girolamo G, Guyer ME, Jin R, et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *Int J Methods Psychiatr Res.* 2006;15(4):167-80.
31. Jacobi F, Wittchen HU, Holting C, Höfler M, Pfister H, Müller N, et al. Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychol Med.* 2004;34(4):597-611.
32. Jacobi F, Höfler M, Siegert J, Mack S, Gerschler A, Scholl L, et al. Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: the Mental Health Module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *Int J Methods Psychiatr Res.* 2014;23(3):304-19.
33. Schmidt CO, Watzke AB, Schulz A, Baumeister SE, Freyberger HJ, Grabe HJ. The lifetime prevalence of mental disorders in north-eastern Germany. What is the influence of earlier mental morbidity on survey participation and prevalence estimates? Results from the SHIP-study. *Psychiatr Prax.* 2013;40(4):192-9.
34. Meyer C, Rumpf HJ, Hapke U, Dilling H, John U. Lifetime prevalence of mental disorders in general adult population. Results of TACOS study. *Nervenarzt.* 2000;71(7):535-42.
35. Asselmann E, Beesdo-Baum K, Hamm A, Schmidt CO, Hertel J, Grabe HJ, et al. Lifetime and 12-month prevalence estimates for mental disorders in northeastern Germany: findings from the Study of Health in Pomerania. *Eur Arch Psychiatry Clin Neurosci.* 2019;269(3):341-50.
36. Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. *The Cochrane database of systematic reviews.* 2005;11(4):CD002792.
37. Gelaye B, Tadesse MG, Williams MA, Fann JR, Vander Stoep A, Andrew Zhou XH. Assessing validity of a depression screening instrument in the absence of a gold standard. *Ann Epidemiol.* 2014;24(7):527-31.



38. Busch MA, Maske UE, Ryl L, Schlack R, Hapke U. Prevalence of depressive symptoms and diagnosed depression among adults in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2013;56(5-6):733-9.
39. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13.
40. Reeves WC, Strine TW, Pratt LA, Thompson W, Ahluwalia I, Dhingra SS, et al. Mental illness surveillance among adults in the United States. *MMWR supplements*. 2011;60(3):1-29.
41. Tomitaka S, Kawasaki Y, Ide K, Akutagawa M, Yamada H, Ono Y, et al. Distributional patterns of item responses and total scores on the PHQ-9 in the general population: data from the National Health and Nutrition Examination Survey. *BMC Psychiatry*. 2018;18(1):108.
42. Shin C, Ko YH, An H, Yoon HK, Han C. Normative data and psychometric properties of the Patient Health Questionnaire-9 in a nationally representative Korean population. *BMC Psychiatry*. 2020;20(1):194.
43. Ślusarska BJ, Nowicki G, Piasecka H, Zarzycka D, Mazur A, Saran T, et al. Validation of the Polish language version of the Patient Health Questionnaire-9 in a population of adults aged 35-64. *Ann Agric Environ Med*. 2019;26(3):420-4.
44. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2006;28(1):71-7.
45. Kocalevent RD, Hinz A, Brähler E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2013;35(5):551-5.
46. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2014;36(5):539-44.
47. Yu X, Tam WW, Wong PT, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. *Compr Psychiatry*. 2012;53(1):95-102.

48. Wang YP, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Revista brasileira de psiquiatria (Sao Paulo, Brazil : 1999)*. 2013;35(4):416-31.
49. Kojima M, Furukawa TA, Takahashi H, Kawai M, Nagaya T, Tokudome S. Cross-cultural validation of the Beck Depression Inventory-II in Japan. *Psychiatry Res*. 2002;110(3):291-9.
50. Segal DL, Coolidge FL, Cahill BS, O'Riley AA. Psychometric properties of the Beck Depression Inventory II (BDI-II) among community-dwelling older adults. *Behav Modif*. 2008;32(1):3-20.
51. Subica AM, Fowler JC, Elhai JD, Frueh BC, Sharp C, Kelly EL, et al. Factor structure and diagnostic validity of the Beck Depression Inventory-II with adult clinical inpatients: comparison to a gold-standard diagnostic interview. *Psychol Assess*. 2014;26(4):1106-15.
52. Beck A. T., Steer R. A., Brown G. K. Beck Depression Inventory (BDI-II), Second Edition. San Antonio, TX: The Psychological Corporation; 1996. 38 p.
53. Desseilles M, Perroud N, Guillaume S, Jaussent I, Genty C, Malafosse A, et al. Is it valid to measure suicidal ideation by depression rating scales? *J Affect Disord*. 2012;136(3):398-404.
54. Simon GE, Rutter CM, Peterson D, Oliver M, Whiteside U, Operskalski B, et al. Does response on the PHQ-9 Depression Questionnaire predict subsequent suicide attempt or suicide death? *Psychiatr Serv*. 2013;64(12):1195-202.
55. Simon GE, Coleman KJ, Rossom RC, Beck A, Oliver M, Johnson E, et al. Risk of suicide attempt and suicide death following completion of the Patient Health Questionnaire depression module in community practice. *J Clin Psychiatry*. 2016;77(2):221-7.
56. Rossom RC, Coleman KJ, Ahmedani BK, Beck A, Johnson E, Oliver M, et al. Suicidal ideation reported on the PHQ9 and risk of suicidal behavior across age groups. *J Affect Disord*. 2017;215:77-84.
57. Louzon SA, Bossarte R, McCarthy JF, Katz IR. Does Suicidal Ideation as Measured by the PHQ-9 Predict Suicide Among VA Patients? *Psychiatr Serv*. 2016;67(5):517-22.
58. Turner AP, Williams RM, Bowen JD, Kivlahan DR, Haselkorn JK. Suicidal ideation in multiple sclerosis. *Arch Phys Med Rehabil*. 2006;87(8):1073-8.

59. Lawrence ST, Willig JH, Crane HM, Ye J, Aban I, Lober W, et al. Routine, self-administered, touch-screen, computer-based suicidal ideation assessment linked to automated response team notification in an HIV primary care setting. *Clin Infect Dis*. 2010;50(8):1165-73.
60. Walker J, Hansen CH, Butcher I, Sharma N, Wall L, Murray G, et al. Thoughts of death and suicide reported by cancer patients who endorsed the "suicidal thoughts" item of the PHQ-9 during routine screening for depression. *Psychosomatics*. 2011;52(5):424-7.
61. Tektonidou MG, Dasgupta A, Ward MM. Suicidal ideation among adults with arthritis: prevalence and subgroups at highest risk. Data from the 2007-2008 National Health and Nutrition Examination Survey. *Arthritis Care Res (Hoboken)*. 2011;63(9):1322-33.
62. Lossnitzer N, Muller-Tasch T, Lowe B, Zugck C, Nelles M, Remppis A, et al. Exploring potential associations of suicidal ideation and ideas of self-harm in patients with congestive heart failure. *Depress Anxiety*. 2009;26(8):764-8.
63. Viguera AC, Milano N, Laurel R, Thompson NR, Griffith SD, Baldessarini RJ, et al. Comparison of Electronic Screening for Suicidal Risk With the Patient Health Questionnaire Item 9 and the Columbia Suicide Severity Rating Scale in an Outpatient Psychiatric Clinic. *Psychosomatics*. 2015;56(5):460-9.
64. Na PJ, Yaramala SR, Kim JA, Kim H, Goes FS, Zandi PP, et al. The PHQ-9 Item 9 based screening for suicide risk: a validation study of the Patient Health Questionnaire (PHQ)-9 Item 9 with the Columbia Suicide Severity Rating Scale (C-SSRS). *J Affect Disord*. 2018;232:34-40.
65. Lichtman JH, Bigger JT, Jr., Blumenthal JA, Frasure-Smith N, Kaufmann PG, Lesperance F, et al. AHA science advisory. Depression and coronary heart disease. Recommendations for screening, referral, and treatment. A science advisory from the American Heart Association Prevention Committee to the Council on Cardiovascular Nursing, Council on Clinical Cardiology, Council on Epidemiology and Prevention, and Interdisciplinary Council on Quality of Care Outcomes Research. Endorsed by the American Psychiatric Association. *Prog Cardiovasc Nurs*. 2009;24(1):19-26.
66. Thompson R, Henkel V, Coyne JC. Suicidal ideation in primary care: ask a vague question, get a confusing answer. *Psychosom Med*. 2004;66(3):455-6.

67. Razykov I, Ziegelstein RC, Whooley MA, Thombs BD. The PHQ-9 versus the PHQ-8 – is item 9 useful for assessing suicide risk in coronary artery disease patients? Data from the Heart and Soul Study. *J Psychosom Res.* 2012;73(3):163-8.
68. Razykov I, Hudson M, Baron M, Thombs BD, Canadian Scleroderma Research G. Utility of the Patient Health Questionnaire-9 to assess suicide risk in patients with systemic sclerosis. *Arthritis Care Res (Hoboken).* 2013;65(5):753-8.
69. Green KL, Brown GK, Jager-Hyman S, Cha J, Steer RA, Beck AT. The Predictive Validity of the Beck Depression Inventory Suicide Item. *J Clin Psychiatry.* 2015;76(12):1683-6.
70. Hoffmann M. Two basic ethical problems of incidental findings in population-based, non-intervening magnetic resonance imaging (MRI) research. *J Eval Clin Pract.* 2013;19(3):427-32.
71. Richter A, Sierocinski E, Singer S, Bülow R, Hackmann C, Chenot JF, et al. The effects of incidental findings from whole-body MRI on the frequency of biopsies and detected malignancies or benign conditions in a general population cohort study. *Eur J Epidemiol.* 2020;35(10):925-35.
72. Booth TC, Waldman AD, Wardlaw JM, Taylor SA, Jackson A. Management of incidental findings during imaging research in "healthy" volunteers: current UK practice. *Br J Radiol.* 2012;85(1009):11-21.
73. Orme NM, Fletcher JG, Siddiki HA, Harmsen WS, O'Byrne MM, Port JD, et al. Incidental findings in imaging research: evaluating incidence, benefit, and burden. *Arch Intern Med.* 2010;170(17):1525-32.
74. Miller FG, Mello MM, Joffe S. Incidental findings in human subjects research: what do investigators owe research participants? *J Law Med Ethics.* 2008;36(2):271-9, 11.
75. Illes J, Kirschen MP, Edwards E, Bandettini P, Cho MK, Ford PJ, et al. Practical approaches to incidental findings in brain imaging research. *Neurology.* 2008;70(5):384-90.
76. Orzechowska A, Zajączkowska M, Talarowska M, Gałecki P. Depression and ways of coping with stress: a preliminary study. *Med Sci Monit.* 2013;19:1050-6.

77. Völzke H, Schössow J, Schmidt CO, Jürgens C, Richter A, Werner A, et al. Cohort Profile Update: The Study of Health in Pomerania (SHIP). *Int J Epidemiol*. 2022;51:e372–83.
78. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1999;56(7):617-26.
79. Wang X, Cheng Z. Cross-Sectional Studies: Strengths, Weaknesses, and Recommendations. *Chest*. 2020;158(1s):S65-s71.
80. Volzke H. Study of Health in Pomerania (SHIP). Concept, design and selected results. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2012;55(6-7):790-4.
81. Schulz A, Schmidt CO, Appel K, Mahler J, Spitzer C, Wingenfeld K, et al. Psychometric functioning, socio-demographic variability of childhood maltreatment in the general population and its effects of depression. *Int J Methods Psychiatr Res*. 2014;23(3):387-400.
82. Wittchen H-U, Pfister H. DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft. Frankfurt: Swets & Zeitlinger; 1997. 261 p.
83. Wittchen HU, Lachner G, Wunderlich U, Pfister H. Test-retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). *Soc Psychiatry Psychiatr Epidemiol*. 1998;33(11):568-78.
84. Lachner G, Wittchen HU, Perkonigg A, Holly A, Schuster P, Wunderlich U, et al. Structure, content and reliability of the Munich-Composite International Diagnostic Interview (M-CIDI) substance use sections. *Eur Addict Res*. 1998;4(1-2):28-41.
85. Wittchen H-U, Höfler M, Gander F, Pfister H, Storz S, Üstün B, et al. Screening for mental disorders: performance of the Composite International Diagnostic – Screener (CID–S). *Int J Methods Psychiatr Res*. 1999;8(2):59-70.
86. Levis B, Benedetti A, Thombs BD. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ*. 2019;365:l1476.
87. Grothe KB, Dutton GR, Jones GN, Bodenlos J, Ancona M, Brantley PJ. Validation of the Beck Depression Inventory-II in a low-income African American sample of medical outpatients. *Psychol Assess*. 2005;17(1):110-4.

88. Kühner C, Bürger C, Keller F, Hautzinger M. Reliability and validity of the Revised Beck Depression Inventory (BDI-II). Results from German samples. *Nervenarzt*. 2007;78(6):651-6.
89. Hedegaard H, Schoenbaum M, Claassen C, Crosby A, Holland K, Proescholdbell S. Issues in Developing a Surveillance Case Definition for Nonfatal Suicide Attempt and Intentional Self-harm Using International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Coded Data. *National health statistics reports*. 2018(108):1-19.
90. Maske UE, Busch MA, Jacobi F, Beesdo-Baum K, Seiffert I, Wittchen HU, et al. Current major depressive syndrome measured with the Patient Health Questionnaire-9 (PHQ-9) and the Composite International Diagnostic Interview (CIDI): results from a cross-sectional population-based study of adults in Germany. *BMC Psychiatry*. 2015;15:77.
91. Jousilahti P, Salomaa V, Kuulasmaa K, Niemelä M, Vartiainen E. Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. *J Epidemiol Community Health*. 2005;59(4):310-5.
92. Wong SY, Mercer SW, Woo J, Leung J. The influence of multi-morbidity and self-reported socio-economic standing on the prevalence of depression in an elderly Hong Kong population. *BMC Public Health*. 2008;8:119.
93. Statistisches Amt Mecklenburg-Vorpommern. *Statistisches Jahrbuch der Landeshauptstadt Schwerin*: Statistisches Amt Mecklenburg-Vorpommern; 2018 [Available from: [https://www.schwerin.de/export/sites/default/.galleries/Dokumente/Stadtporraet/Statistik-Jahrbuch-2018\\_r.pdf](https://www.schwerin.de/export/sites/default/.galleries/Dokumente/Stadtporraet/Statistik-Jahrbuch-2018_r.pdf) (last accessed: 13.01.2021).
94. Harald K, Salomaa V, Jousilahti P, Koskinen S, Vartiainen E. Non-participation and mortality in different socioeconomic groups: the FINRISK population surveys in 1972-92. *J Epidemiol Community Health*. 2007;61(5):449-54.
95. Drivsholm T, Eplöv LF, Davidsen M, Jørgensen T, Ibsen H, Hollnagel H, et al. Representativeness in population-based studies: a detailed description of non-response in a Danish cohort study. *Scandinavian journal of public health*. 2006;34(6):623-31.

96. Larsen SB, Dalton SO, Schuz J, Christensen J, Overvad K, Tjonneland A, et al. Mortality among participants and non-participants in a prospective cohort study. *Eur J Epidemiol.* 2012;27(11):837-45.
97. Statistisches Bundesamt. Anzahl der Suizide in Deutschland nach Bundesländern in den Jahren 2010 bis 2020: Statistisches Bundesamt; 2022 [Available from: <https://de.statista.com/statistik/daten/studie/218237/umfrage/anzahl-der-suizide-in-deutschland-nach-bundeslaendern/> (last accessed: 05.06.2022).
98. Vennemann MMT. Unterschätzte Suizidraten durch unterschiedliche Erfassung in Gesundheitsämtern. *Dtsch Arztebl International.* 2006;103(18):1222-6.
99. Rockett IR, Hobbs GR, Wu D, Jia H, Nolte KB, Smith GS, et al. Variable Classification of Drug-Intoxication Suicides across US States: A Partial Artifact of Forensics? *PLoS One.* 2015;10(8):e0135296.
100. Skinner R, McFaull S, Rhodes AE, Bowes M, Rockett IRH. Suicide in Canada: Is Poisoning Misclassification an Issue? *The Canadian Journal of Psychiatry.* 2016;61(7):405-12.
101. Stone DM, Holland KM, Bartholow B, J EL, LiKamWa McIntosh W, Trudeau A, et al. Deciphering Suicide and Other Manners of Death Associated with Drug Intoxication: A Centers for Disease Control and Prevention Consultation Meeting Summary. *Am J Public Health.* 2017;107(8):1233-9.
102. Kroenke K, Spitzer RL, Williams JB, Lowe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *Gen Hosp Psychiatry.* 2010;32(4):345-59.
103. Hunt M, Auriemma J, Cashaw AC. Self-report bias and underreporting of depression on the BDI-II. *J Pers Assess.* 2003;80(1):26-30.
104. Hill NTM, Robinson J, Pirkis J, Andriessen K, Krysinska K, Payne A, et al. Association of suicidal behavior with exposure to suicide and suicide attempt: A systematic review and multilevel meta-analysis. *PLoS Med.* 2020;17(3):e1003074.
105. Nie J, O'Neil A, Liao B, Lu C, Aune D, Wang Y. Risk factors for completed suicide in the general population: A prospective cohort study of 242, 952 people. *J Affect Disord.* 2021;282:707-11.

106. Batty GD, Kivimäki M, Bell S, Gale CR, Shipley M, Whitley E, et al. Psychosocial characteristics as potential predictors of suicide in adults: an overview of the evidence with new results from prospective cohort studies. *Translational psychiatry*. 2018;8(1):22.



## Appendix A

**Table A.1:** Age-specific mortality in Mecklenburg-Western Pomerania (reporting year 2017)

<b>Age</b>	<b>Mortality<sup>a</sup></b>
30<35	0.5
35<40	0.8
40<45	1.8
45<50	2.8
50<55	6.6
55<60	7.5
60<65	10.5
65<70	19.3
70<75	15.9
75<80	30.7
80<85	52.1
85<90	93.2

<sup>a</sup>per 1.000 inhabitants at the same age (93).

## Appendix B (Instruments)

**Table B.1:** Composite International Diagnostic-Screener (CID-S) in SHIP-START-0

Symptom	Question
Daily sadness	Have you ever experienced two weeks or more when you felt sad, blue or depressed almost every day?
Daily loss of energy	Did you ever experience loss of energy, or feeling tired almost every day over a period of two weeks or longer?

**Table B.2:** Composite International Diagnostic-Screener (CID-S) in SHIP-START-1, SHIP-START-2, SHIP-START-3

Symptom	Question
Daily sadness	Did you have feelings of sadness or despondency daily over a period of more than 2 weeks during the last 12 months?
Daily loss of energy	Did you have a loss of interest, tiredness or a loss of energy daily over a period of more than 2 weeks during the last 12 months?

**Table B.3:** Patient Health Questionnaire 9 (PHQ-9)

Item	Question
1	Over the last 2 weeks, how often have you been bothered by little interest or pleasure in doing things?
2	Over the last 2 weeks, how often have you been bothered by Feeling down, depressed, or hopeless?
3	Over the last 2 weeks, how often have you been bothered by trouble falling or staying asleep, or sleeping too much?
4	Over the last 2 weeks, how often have you been bothered by feeling tired or having little energy?
5	Over the last 2 weeks, how often have you been bothered by poor appetite or overeating?
6	Over the last 2 weeks, how often have you been bothered by feeling bad about yourself — or that you are a failure or have let yourself or your family down?
7	Over the last 2 weeks, how often have you been bothered by trouble concentrating on things, such as reading the newspaper or watching television?
8	Over the last 2 weeks, how often have you been bothered by moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual?
9	Over the last 2 weeks, how often have you been bothered by thoughts that you would be better off dead or of hurting yourself in some way?

**Table B.4:** Beck Depression Inventory-II (BDI-II)

<b>Item</b>	<b>Answer option</b>
1. Sadness	0. I do not feel sad. 1. I feel sad much of the time. 2. I am sad all the time. 3. I am so sad or unhappy that I can't stand it.
2. Pessimism	0. I am not discouraged about my future. 1. I feel more discouraged about my future than I used to. 2. I do not expect things to work out for me. 3. I feel my future is hopeless and will only get worse.
3. Past Failure	0. I do not feel like a failure. 1. I have failed more than I should have. 2. As I look back, I see a lot of failures. 3. I feel I am a total failure as a person.
4. Loss of Pleasure	0. I get as much pleasure as I ever did from the things I enjoy. 1. I don't enjoy things as much as I used to. 2. I get very little pleasure from the things I used to enjoy. 3. I can't get any pleasure from the things I used to enjoy.
5. Guilty Feelings	0. I don't feel particularly guilty. 1. I feel guilty over many things I have done or should have done. 2. I feel quite guilty most of the time. 3. I feel guilty all of the time.
6. Punishment Feelings	0. I don't feel I am being punished. 1. I feel I may be punished. 2. I expect to be punished. 3. I feel I am being punished.
7. Self-Dislike	0. I feel the same about myself as ever. 1. I have lost confidence in myself. 2. I am disappointed in myself. 3. I dislike myself.
8. Self-Criticalness	0. I don't criticize or blame myself more than usual. 1. I am more critical of myself than I used to be. 2. I criticize myself for all of my faults. 3. I blame myself for everything bad that happens.
9. Suicidal Thoughts or Wishes	0. I don't have any thoughts of killing myself. 1. I have thoughts of killing myself, but I would not carry them out. 2. I would like to kill myself. 3. I would kill myself if I had the chance.
10. Crying	0. I don't cry anymore than I used to. 1. I cry more than I used to. 2. I cry over every little thing. 3. I feel like crying, but I can't.
11. Agitation	0. I am no more restless or wound up than usual. 1. I feel more restless or wound up than usual. 2. I am so restless or agitated, it's hard to stay still. 3. I am so restless or agitated that I have to keep moving or doing something.
12. Loss of Interest	0. I have not lost interest in other people or activities. 1. I am less interested in other people or things than before. 2. I have lost most of my interest in other people or things. 3. It's hard to get interested in anything.
13. Indecisiveness	0. I make decisions about as well as ever. 1. I find it more difficult to make decisions than usual. 2. I have much greater difficulty in making decisions than I used to. 3. I have trouble making any decisions.
14. Worthlessness	0. I do not feel I am worthless. 1. I don't consider myself as worthwhile and useful as I used to. 2. I feel more worthless as compared to others. 3. I feel utterly worthless.

**Table B.4:** Beck Depression Inventory-II (BDI-II), Continued

<b>Item</b>	<b>Answer option</b>
15. Loss of Energy	0. I have as much energy as ever. 1. I have less energy than I used to have. 2. I don't have enough energy to do very much. 3. I don't have enough energy to do anything.
16. Changes in Sleeping Pattern	0. I have not experienced any change in my sleeping. 1a I sleep somewhat more than usual. 1b I sleep somewhat less than usual. 2a I sleep a lot more than usual. 2b I sleep a lot less than usual. 3a I sleep most of the day. 3b I wake up 1-2 hours early and can't get back to sleep.
17. Irritability	0. I am not more irritable than usual. 1. I am more irritable than usual. 2. I am much more irritable than usual. 3. I am irritable all the time.
18. Changes in Appetite	0. I have not experienced any change in my appetite. 1a My appetite is somewhat less than usual. 1b My appetite is somewhat greater than usual. 2a My appetite is much less than before. 2b My appetite is much greater than usual. 3a I have no appetite at all. 3b I crave food all the time.
19. Concentration Difficulty	0. I can concentrate as well as ever. 1. I can't concentrate as well as usual. 2. It's hard to keep my mind on anything for very long. 3. I find I can't concentrate on anything.
20. Tiredness or Fatigue	0. I am no more tired or fatigued than usual. 1. I get more tired or fatigued more easily than usual. 2. I am too tired or fatigued to do a lot of the things I used to do. 3. I am too tired or fatigued to do most of the things I used to do.
21. Loss of Interest in Sex	0. I have not noticed any recent change in my interest in sex. 1. I am less interested in sex than I used to be. 2. I am much less interested in sex now. 3. I have lost interest in sex completely.

## List of Abbreviations

AHA	American Heart Association
BDI	Beck Depression Inventory
CIDI	Composite International Diagnostic Interview
CID-S	Composite International Diagnostic-Screener
DEGS	German Health Interview and Examination Survey
DSM	Diagnostic and Statistical Manual of Mental Disorders
ICD	International Classification of Diseases
LEGENDE	Life-Events and Gene-Environment Interaction in Depression
M-CIDI	Munich-Composite International Diagnostic Interview
NHANES	National Health and Nutrition Examination Survey
NPV	Negative predictive value
PHQ	Patient Health Questionnaire
PPV	Positive predictive value
SHIP	Study of Health in Pomerania
VHA	Veterans Health Administration
WHO	World Health Organization

## List of Tables

<b>Table 1:</b> Suicidal ideation in population-based studies	2
<b>Table 2:</b> Overview of instruments used for the measurement of depression and depressive symptoms	9
<b>Table 3:</b> SHIP-START-0 and SHIP-TREND-0 baseline characteristics	14
<b>Table 4:</b> Depression (by M-CIDI) and depressive symptoms (by CID-S, PHQ-9 and BDI-II) by gender	15
<b>Table 5:</b> Suicidal ideation (by single items of PHQ-9 and BDI-II) by gender	17
<b>Table 6:</b> Mortality and potential suicide by time	18
<b>Table 7:</b> Potential suicide after indications of depression (by M-CIDI) and depressive symptoms (by CID-S, PHQ-9, BDI-II) by time	19
<b>Table 8:</b> Potential suicide and other causes of death after indications of suicidal ideation (by single items of PHQ-9 and BDI-II) by time	20

## **Eidesstattliche Erklärung**

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Die Dissertation ist bisher keiner anderen Fakultät, keiner anderen wissenschaftlichen Einrichtung vorgelegt worden.

Ich erkläre, dass ich bisher kein Promotionsverfahren erfolglos beendet habe und dass eine Aberkennung eines bereits erworbenen Doktorgrades nicht vorliegt.

---

Datum

---

Unterschrift