Negative association of concomitant physical symptoms with the course of major depressive disorder: A systematic review


Abstract

Objective: The prognosis of depression greatly varies among patients, and the physical symptoms that often accompany depression may predict treatment resistance and a worse outcome. If so, this may have important clinical implications. The aim of this systematic review was to explore the association of concomitant physical symptoms with the outcome of Major Depressive Disorder (MDD).

Methods: Systematic review. Medline, Psychinfo, and the Cochrane Library were searched for prospective, cross-sectional, and retrospective studies, and also for open-label trials and randomised controlled trials. The risk of bias assessment and data-extraction were performed in duplicate. A qualitative best-evidence synthesis was performed, based on the number of studies reporting on the association between physical symptoms and the course of MDD, the consistency of the results, and the methodological quality. The findings were reported according to the PRISMA guidelines.

Results: Nine studies met the inclusion criteria. Although the design, outcome measures, and data presentation varied too much to make statistical pooling possible, the best evidence synthesis resulted in strong, consistent evidence for a negative association between physical symptoms and the course of MDD.

Conclusion: This systematic review shows a negative association of concomitant physical symptoms with the course of MDD. The effect might be considerable, but the number of studies addressing this topic is small and there was a wide variation in the study designs and outcome measures. More research is needed.

Keywords: depression, comorbidity, effect modifiers (epidemiology), somatisation, systematic review, physical symptoms
Introduction

The Global Burden of Disease Study, carried out by the WHO, reports that major depressive disorder (MDD) is expected to be the second leading cause of disability-adjusted life-years in 2020.\textsuperscript{6} Both psychological and pharmacological interventions are effective in the treatment of MDD, but like in other fields of medicine, results from efficacy trials where it is possible to control certain conditions, do not always produce the same results in effectiveness trials in real world settings.\textsuperscript{72} Results from the STAR*D study, for instance, show that remission rates in response to treatment with an SSRI are as low as 26.6\% in primary care.\textsuperscript{147} These rates improve after subsequent treatment steps, but remission is not easily achieved.\textsuperscript{148} This might lead to unnecessary suffering and high costs for society.\textsuperscript{149;150}

MDD may be difficult to treat in everyday practice, as a result of concomitant physical symptoms interfering with the course, or because depressed patients who experience many physical symptoms may be less motivated to undergo treatment. This is an important subject for study, in view of the frequent co-occurrence of physical symptoms and MDD. Studies in primary care, for instance, have found that up to 70\% of MDD patients only report physical symptoms when first presenting to a general practitioner (GP).\textsuperscript{25} Moreover, a recent study\textsuperscript{29} reported that it is 4.43 times more likely for a depressed patient to have a somatoform disorder than for a patient who is not depressed (Confidence Interval: 2.73-7.19).

Examples of symptoms that often co-occur with MDD are pain, fatigue, disturbed sleep, indigestion, dizziness, and fainting.\textsuperscript{151;152} The more of these symptoms a patient experiences, the greater the probability that he or she is also suffering from a depressive disorder.\textsuperscript{151;152} The importance of physical symptoms in patients suffering from MDD will also be recognised by clinicians, many of whom find patients with physical co-morbidity difficult to treat. A recent systematic review provided evidence that pain predicts a longer time to remission in patients suffering from MDD.\textsuperscript{153}

To our knowledge, evidence with regard to the prognostic effect of the wider spectrum of concomitant physical symptoms -not restricted to pain- on the course of MDD has not yet been reviewed. Identifying factors that may predict treatment resistance or poor outcome are important, both for the education of patients and the selection of appropriate treatment. We therefore conducted a systematic review of studies that assessed the associations between physical symptoms and the prognosis of MDD. The association in terms of prognostic value in cohort studies, as well as the modification of treatment effect in trials can provide valuable information about this topic. Both types of studies were therefore included.
Methods

Information sources and eligibility criteria

The search was performed in Medline, Psychinfo, and the database of the Cochrane Collaboration. The search focused on retrospective and prospective designs, as well as on open-label trials and randomised controlled trials (RCTs). Studies had to meet the following criteria in order to be included:

1. At baseline an assessment had to be made to determine whether patients were suffering from MDD. There was no restriction as to how MDD was diagnosed. At follow-up an assessment had to be made about the outcome of the disorder. This could include clinical outcomes such as symptom reduction on a scale measuring depressive symptoms, response in terms of a decrease in a certain percentage on a scale measuring depressive symptoms, and time to such a response. It could also include process measures such as treatment adherence.

2. An assessment of physical symptoms had to be included at baseline. Since this review focused on a wider spectrum of physical symptoms than just one individual symptom, such as pain, special attention was paid to the way in which these symptoms were identified. The assessment of physical symptoms at baseline could not be limited to just one symptom. Furthermore, the studies had to include either:
   (a) a validated (sub-)scale to measure physical symptoms
   (b) information about physical symptoms from medical records
   (c) an assessment of physical symptoms by a medical practitioner.

With regard to the sub-criterion (a) (validated measures), we decided that the anxiety/somatization sub-scale of the Hamilton Depression Inventory 17-item version (HAMD-17)\textsuperscript{154} could not be used as a valid measure of physical symptoms for the purpose of our review. This sub-scale consists of six items that reflect on both psychological and somatic symptoms of depression and anxiety. Therefore, studies had to use some other instrument to assess physical symptoms.

3. Studies had to report sufficient information on the association between physical symptoms and the outcome or course of MDD, expressed as Odds Ratios (OR), Relative Risks (RR), Beta-coefficients ($\beta$), Hazard Ratios (HR), effect sizes, or mean differences.
Search

The exact keywords and Mesh-terms are available upon request. Briefly, we combined multiple keywords and medical subject headings (Mesh-terms) for MDD with multiple keywords regarding information on the outcome or course of the disorder and keywords referring to the design of the study. Finally, 102 symptoms and symptom-related keywords were added to the search, ranging from pain and fatigue to somatoform disorders or somatisation, and including symptoms such as headache, nausea, and gas. Our search strategy is described in more detail in the Appendix.

The following searches were conducted:

- Medline was searched from its onset to November 24th 2008. There was no restriction as to the publication date of a study. The searches in Medline were periodically updated after the 24th of November 2008, but no additional papers were identified before submission of this document. There was no language restriction included in the search.
- Psychinfo was searched until March 25th 2009. Again, there was no restriction with regard to publication date or language.
- The Cochrane Library database was searched until March 25th 2009, again with no restrictions with regard to publication date or language.

Study selection, risk of bias assessment, and data-collection process

All titles and abstracts (N=2327) were read by one of the authors (KH). Full-text publications of studies that were potentially eligible for inclusion (N=82) were read by two reviewers (K.H. and C.F.C.). They both checked the criteria for eligibility. Consensus was reached in all cases. In this stage the quality of the papers was not assessed. Nine papers met the eligibility criteria. These were included in the review, and were submitted to full quality assessment, performed independently by two assessors (K.H. and C.F.C.) who were unaware of each other's assessments. The studies were assessed according to a list of predefined criteria described in a systematic review.\(^{155}\) This checklist is based on previously used validity checklists\(^{156}\), theoretical considerations, and methodological aspects of prognostic research.\(^{157,158}\) Any disagreements among the reviewers were resolved by asking a third independent assessor (HvM) to assess the study on the criterion that caused disagreement. In all cases consensus was reached. The list of criteria is described in the first column of Table 2. A study was considered to be of high quality if 60% or more of the quality criteria were met (10 or more criteria out of 16).

Data were extracted on changes in the severity of depressive symptoms and the association of this outcome with physical symptoms. The type of association varied among the studies, and could be presented as ORs, regression-coefficients, correlation coefficients or mean changes. Data were also extracted on the design of the study, the setting, the characteristics of the study population, the treatment of depression, and the duration of the follow-up.
Data-extraction was performed independently, and in duplicate, by two members of the research group (KH and CFC).

**Synthesis of results**

The methods used to describe the outcome of MDD and the methods used to present the association with physical symptoms varied widely in the studies that were included in this review. Outcome was presented as: time to a clinically relevant response to treatment, probability of a clinically relevant response, or mean change on a scale measuring depressive symptoms. Also presented were the association between physical symptoms and adherence to pharmacotherapy, or adverse effects of the treatment.

The outcome measures essentially reflected two types of association between physical symptoms and MDD: they either described the association between physical symptoms and the outcome of MDD (type A) or they described the association between physical symptoms and adherence to pharmacotherapy for MDD (type B). These two types of associations were analyzed separately.

The way in which the studies operationally defined association varied considerably. This variation in outcome assessment and data-presentation precluded an overall meta-analysis, because this would not provide meaningful estimates of effect for one clinically relevant outcome measure. Therefore, we performed a qualitative analysis and presented a best-evidence synthesis. The number of studies reporting on the association between physical symptoms and the course of MDD, the consistency of the results, and the methodological quality were taken into account. A priori, findings were considered to be consistent if 75% or more of the studies reporting on a factor showed the same direction of the association. A study was considered to be of high quality if it met 60% or more of the quality criteria described in the first column of Table 2.

The definition of a level of evidence for the direction of association was based on earlier methods used to summarise the findings of observational research. The level of evidence could have one of the following values:

- **Strong**: consistent and statistically significant associations in at least two high-quality cohorts.
- **Moderate**: consistent and statistically significant association in one high-quality cohort and at least one low-quality cohort.
- **Weak**: statistically significant association in one high-quality cohort or consistent and statistically significant in at least three low-quality cohorts.
- **Inconclusive**: association in less than three low-quality cohorts, regardless of statistical significance.
- **Inconsistent**: inconsistent findings, irrespective of study quality.
- **Insufficient**: only one study presenting non-statistical significant association, irrespective of quality.

---

Results

**Study selection and characteristics**

The searches in Medline, Psychinfo, and the Cochrane Library yielded 2327 hits. All of the abstracts were written in a language that could be interpreted by the authors. Nine studies met the inclusion criteria.

The results of the searches are shown in a flowchart (Figure 1) that was designed according to the PRISMA guidelines. The characteristics of the studies that were included are presented in Table 1.

**Figure 1  Study selection**

2327 records identified through database screening:
- Medline: 1129
- Psychinfo: 1198
- Cochrane Library: 0

2245 records screened
2245 records after duplicates were removed
82 full-text articles assessed for eligibility
2163 records excluded
- MDD was not an outcome
- Physical symptoms were not included at baseline
- Design of the study other than trial + cohort (for instance comments)
- No abstract
9 studies included in qualitative synthesis
73 full-text articles excluded on closer examination because physical symptoms were not operationally defined in a satisfactory way (HAMD17 or only pain or fatigue and no focus on other symptoms).
4 additional records identified through other sources
Results

Study selection and characteristics

The searches in Medline, Psychinfo, and the Cochrane Library yielded 2327 hits. All of the abstracts were written in a language that could be interpreted by the authors. Nine studies met the inclusion criteria.

The results of the searches are shown in a flowchart (Figure 1) that was designed according to the PRISMA guidelines. The characteristics of the studies that were included are presented in Table 1.

Figure 1 Study selection

2327 records identified through database screening:
- Medline: 1129
- Psychinfo: 1198
- Cochrane Library: 0

2245 records after duplicates were removed

2245 records screened

82 full-text articles assessed for eligibility

9 studies included in qualitative synthesis

2163 records excluded
Records were excluded for the following reasons:
- MDD was not an outcome
- Physical symptoms were not included at baseline
- Design of the study other than trial + cohort (for instance comments)
- No abstract

73 full-text articles excluded on closer examination because physical symptoms were not operationally defined in a satisfactory way (HAMD17 or only pain or fatigue and no focus on other symptoms).
<table>
<thead>
<tr>
<th>Study/Country/ Setting</th>
<th>N</th>
<th>Drop out perc.*</th>
<th>Follow-up period</th>
<th>Type of intervention</th>
<th>Definition of depression</th>
<th>Definition of physical symptoms</th>
<th>Description of effect</th>
<th>Effects size (CI if sufficient data was available)</th>
<th>Type of association/ Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karp et al. United States, Psychiatric clinic</td>
<td>230</td>
<td>8.3%</td>
<td>25 weeks</td>
<td>Imipramine hydrochloride 150-300 mg. + IPT</td>
<td>HRSD≥15</td>
<td>SCL90-R somatic scale</td>
<td>Median time to remission on HAMD-17 (score≤7) longer in group reporting somatic symptoms (19.1 weeks versus 12.9 weeks)</td>
<td>HR: .80</td>
<td>A (69%; high quality)</td>
</tr>
<tr>
<td>Papakostas et al. United States, Outpatients in a general hospital</td>
<td>87</td>
<td>15.6%</td>
<td>8 weeks</td>
<td>Fluoxetine 20 mg.</td>
<td>SCID-P</td>
<td>SQ-SS</td>
<td>Somatic symptom score not significantly related to time to response (50% decrease on HAMD-17), but greater number of somatic symptoms does predict greater time to onset of response (difference in weeks not specified)</td>
<td>β-coefficient=.09 (Se=.039; P=.02) ** (CI=.014-.166)</td>
<td>A (56%; lower quality)</td>
</tr>
<tr>
<td>Papakostas et al. United States, Outpatients</td>
<td>40</td>
<td>22.5%</td>
<td>6 weeks</td>
<td>Nortriptyline 25 mg.-100 mg.</td>
<td>SCID-III-R</td>
<td>SQ-SS</td>
<td>Lower probability of response (50% decrease on HAMD-17)</td>
<td>OR=2.29 (CI=1.13-5.0; P=.02) ***</td>
<td>A (50%; lower quality)</td>
</tr>
<tr>
<td>Papakostas et al. United States, Hospital based academic sites</td>
<td>570</td>
<td>35.8%</td>
<td>12 weeks</td>
<td>Fluoxetine 40 mg. Minimum</td>
<td>SCID-I/P</td>
<td>SCL90-R somatic scale</td>
<td>Lower probability of response (50% decrease on HAMD-17)</td>
<td>OR=.96**** (CI=.94-.98; P&lt;.01)</td>
<td>A (75%; high quality)</td>
</tr>
<tr>
<td>Keeley et al. United States, Primary care</td>
<td>200</td>
<td>11.1%</td>
<td>6 months</td>
<td>Enhanced care</td>
<td>Patients reporting 5 out of 9 DSM-IV criteria of MDD</td>
<td>Physical symptoms extracted from the medical records</td>
<td>Effect of enhanced care (EC) vs care as usual (CAU) was only significant for psychological presenters (P=.04). Psych. pres. in the EC group improved 50.9% after 6 months. Phys. pres. improved 28.7% on mCES-D</td>
<td>A (87%; high quality)</td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Study/Country/Setting</th>
<th>N</th>
<th>Percent drop-out</th>
<th>Follow-up period</th>
<th>Type of intervention</th>
<th>Definition of depression</th>
<th>Definition of physical symptoms</th>
<th>Description of effect</th>
<th>Effects size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoencamp et al.</td>
<td>119</td>
<td>26.1%</td>
<td>18 weeks</td>
<td>Three-phase sequential medication strategy with Maprotiline, Maprotiline + Lithium, Broforamine</td>
<td>HRSD≥14</td>
<td>SCL-90-R somatic scale</td>
<td>Percentage change in HRSD-score</td>
<td>β-coefficient=-.24 (P=.02)</td>
</tr>
<tr>
<td>Papakostas et al.</td>
<td>170</td>
<td>15.6%</td>
<td>8 weeks</td>
<td>Fluoxetine 20 mg.</td>
<td>SCID-P</td>
<td>SQ-SS</td>
<td>Greater degree of somatic symptoms at baseline during at least one TRAE reported as moderate or severe (No mention of severity threshold)</td>
<td>(65%; lower quality)</td>
</tr>
<tr>
<td>Agosti et al.</td>
<td>940</td>
<td>23%</td>
<td>6 weeks</td>
<td>Imipramine, Phenelzine, L-deprenyl, Mianserin, Desipramine</td>
<td>RDC for major depression or DSM-III criteria</td>
<td>SCL90-R somatic scale</td>
<td>Patients who discontinued medication due to adverse reactions had a higher somatic score at baseline than non-dropouts (1.3; SD=.67)</td>
<td>(CI=1.94-2.65)**</td>
</tr>
<tr>
<td>Rollman et al.</td>
<td>91</td>
<td>17.6%</td>
<td>6 weeks</td>
<td>Nortryptiline 25mg, increased until stable serum level of 50-150 ng/mL</td>
<td>CES-D ≥22</td>
<td>SSC</td>
<td>Non-adhering patients had a higher physical score (5.8; SD=4.0) than completers (4.2; SD=2.9)</td>
<td>(CI=4.23-7.37)**</td>
</tr>
</tbody>
</table>

*In most cases the drop-out percentage and follow-up period were derived from the data in the manuscript or from an older study if the manuscript was based on a secondary data analysis. Most articles did not directly state the specific data. ** CI =Confidence interval. *** Greater average decrease in probability of clinical response with each point increase in symptom score on the SCL-90-R somatic sub-scale (i.e. from 0 to 1, 1 to 2, etc.) assuming baseline HAMD-17 scores are identical. 

**Abbreviations**

- CAU=Care as Usual
- CES-D= Center for Epidemiologic Studies-Depression
- CI= Confidence Interval
- EC=Enhanced Care
- HAMD= Hamilton Rating Scale for Depression
- HR= Hazard Ratio
- HRSD= Hamilton Rating Scale for Depression
- IPT= Interpersonal Psychotherapy
- mCES-D= Modified Center for Epidemiological Studies Depression scale
- MDD= Major Depressive Disorder
- N=Number
- OR=Odds Ratio
- Perc.=Percentage
- Phys. Pres.=Physical Presenters
- Psych. Pres.=Psychological Presenters
- RDC= Research Diagnostic Criteria for Depression [29]
- RR= Relative Risk
- RSDS= Raskin Severity of Depression Scale
- SCID-I/P= Structured Clinical Interview for DSM-IV Axis I disorders
- SQ-SS= Symptom Questionnaire Somatic Subscale
- SSC= Somatic Symptoms Checklist
- TRAE= Treatment related Adverse Effect
- TypeA association= Physical symptoms and MDD outcome
- TypeB association= Physical symptoms and adherence to pharmacotherapy
- US= United States
- WSAS= Work and Social Adjustment Scale

Chapter 5
Table 2  Results of quality assessment (risk of bias assessment)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Karp et al.40</th>
<th>Papakostas et al.162</th>
<th>Papakostas et al.163</th>
<th>Papakostas et al.164</th>
<th>Keeley et al.165</th>
<th>Hoencamp et al.166</th>
<th>Papakostas et al.167</th>
<th>Agosti et al.168</th>
<th>Rollman et al.169</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Inception cohort (positive if interval between diagnosis of depression and baseline assessment was 6 weeks or less)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>B. Description of study population (criteria should be formulated for at least age, gender and setting)</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>C. Definition of depression (depression should be diagnosed using structured, validated instruments)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D. N≥100</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Response/Representativeness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. Response rate in target population (percentage of possibly eligible patients that could be assessed for inclusion)≥ 75%</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>F. Information about non-responders vs. responders in target population (positive if information was presented about patient/disease characteristics of responders and non-responders)</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>?</td>
</tr>
<tr>
<td>Follow-up (extent and length)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G. Prospective data-collection</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>H. Follow-up of at least 6 months</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>?</td>
</tr>
<tr>
<td>I. Drop-outs /loss to follow-up &lt; 20%</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>J. Information about completers vs. loss to follow-up/drop-outs (patient/disease characteristics such as age, gender and other potential prognostic predictors)</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Study →</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K. Description of treatment clearly</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>L. Standardised assessment of depression outcome</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Prognostic factors/Physical symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. Standardised assessment of physical symptoms at baseline (symptoms should be measured with a structured and validated instrument and should not be limited to only pain, fatigue or vital signs of Major Depressive Disorder)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Data presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. Frequencies of most important outcome measures (either frequency, percentage, mean, median → Inter-quartile Range)</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>O. Frequencies of physical symptom-scores presented (either frequency, percentage, mean, median → Inter-quartile Range)</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>P. Influence of physical symptoms presented</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Total score per study →</td>
<td>69%</td>
<td>56%</td>
<td>50%</td>
<td>75%</td>
<td>87%</td>
<td>56%</td>
<td>62%</td>
<td>81%</td>
<td>56%</td>
</tr>
</tbody>
</table>
Initial agreement between the assessors was reached in 69% of all cases. All differences in opinion were solved by discussion, and consensus was reached in all cases.

The mean quality score was 66%, ranging from 50% to 87% (on average 10.6 out of 16 criteria were met). Five studies had a high quality score (>60%). The criteria that were most frequently not met were H (follow-up of at least six months) and J (information about completers vs. loss to follow-up/drop outs).

**Synthesis of results**

The results are summarised in Table 1. The association between physical symptoms and the outcomes of the nine studies that were included could be combined to form two types of association: the overall association between physical symptoms and the course of MDD (type A) and the association between physical symptoms and adherence to pharmacotherapy and treatment-related adverse effects (type B). The last column of Table 1 refers to this type of association. An overall level of evidence was determined for both types, and is described in the text below.

Furthermore, a level of evidence was determined for the sub-types that were used to determine the overall level of evidence, namely the association of physical symptoms with respectively: time to a clinically relevant response to treatment, probability of a clinically relevant response, or mean change on a scale measuring depressive symptoms. These levels of evidence are also described below.

**A. Association between physical symptoms and course of MDD: overall**

**Findings**

Six studies reported on the association between physical symptoms and the course of MDD. All six studies reported a negative influence of these symptoms, and two studies\textsuperscript{164,165} had a high score for quality in our review. Both reported statistically significant associations. Based on the criteria for determining the level of evidence, the conclusion is that there is strong evidence for a negative overall association of physical symptoms with the course of MDD.

As mentioned above, the overall association between physical symptoms and course of MDD was composed of three sub-types. The level of evidence for each of these sub-types is described below, as well as the strengths of these associations in the individual studies. Two studies\textsuperscript{40,162} reported on the influence of physical symptoms on the time to response to treatment, with results pointing in the same direction. A high score on a scale measuring physical symptoms at baseline predicted a 6.2-week longer time to treatment response in one study\textsuperscript{40} and a longer time to onset of response in the other.\textsuperscript{162} The duration was not specified in the latter study. The level of evidence for the association in terms of time to response, is inconclusive, since one of the studies reported non-significant results\textsuperscript{40} and the other study\textsuperscript{162} had a low score for validity.

Two other studies\textsuperscript{163,164} reported on the influence of physical symptoms on the probability of response to treatment. Both demonstrated that a high score on a scale measuring physical symptoms at baseline predicts a lower probability of a good response. One of the studies\textsuperscript{163} reported an OR of 2.29 on an unfavourable course for patients with physical symptoms at baseline. The other study\textsuperscript{164} reported a protective OR of a low score on the SCL-90-R somatic scale.\textsuperscript{170} This OR is a .96 per point decrease on
this scale. Therefore, a decrease, for instance, of 10 points on this scale (ranging between 12 and 60) at baseline already implies a protective OR of .61 on a favourable response. Since the results of both studies are significant, but only one study had a high score for quality\textsuperscript{164}, the level of evidence is moderate for this association.

Finally, two studies\textsuperscript{165;166} described a negative and significant association between physical symptoms and severity of depression at follow-up. One of the studies\textsuperscript{165} reported that for patients with a physical presentation of MDD at baseline the decrease on a scale measuring the severity of MDD was almost twice as much as for than patients with a psychological presentation. The other study\textsuperscript{166} reported a Beta-coefficient of -.24, implying that a patient who scores, for instance, 10 points higher on the somatic scale of the SCL-90-R could expect a 2.4-point less decrease on the HAMD-17. The level of evidence for the association between physical symptoms and severity of depression is moderate, since one of the studies had a high score for quality and one study had a lower score for quality.

**B. Association between physical symptoms and adherence to pharmacotherapy and treatment-related adverse effects**

Three studies\textsuperscript{167-169} reported on the association between physical symptoms and adherence to pharmacotherapy or treatment-related adverse effects. Two studies\textsuperscript{168;169} demonstrated a relationship between discontinuation of medication and physical symptoms at baseline. Only one of these studies had a high score for quality.\textsuperscript{168} Since both studies reported statistically significant results, the conclusion is that there is moderate evidence for a negative association between physical symptoms and adherence to pharmacotherapy.

One study\textsuperscript{167} found that there was no correlation between somatic scores on the SQ-SS and number of Treatment-related Adverse Effects (TRAЕ), but that a greater degree of somatic symptoms at baseline predicted emergence of at least one TRAE reported as moderate or severe. Since this is the only high quality study reporting on this relationship, the evidence remains weak.
Discussion

Summary of evidence

All studies included in this systematic review indicated a negative effect of physical symptoms on the prognosis of MDD. Based on the consistency of this finding across the available studies, the strength of the associations found, and the quality of the studies, the overall level of evidence for this association appears to be strong. The overall association of physical symptoms was sub-divided in two types of association: type A, which refers to the association between physical symptoms and the outcome of MDD, and type B, which refers to the association between physical symptoms and adherence to pharmacotherapy for MDD. The evidence for type A association appeared to be strong, and the evidence for type B association appeared to be moderate. The findings for both types of association, and especially the strong evidence for the overall association of physical symptoms with a poorer prognosis of MDD, are important, given the high co-morbidity of physical symptoms with MDD that is reported in the scientific literature.25,29

Ideally one would like to pool the results of several studies, reporting the size of the effect that concomitant physical symptoms have on the prognosis of depression, but given the small number of studies available and the variety of methods and measures employed, calculating a meaningful overall effect-size was not feasible. However, the available studies do warrant the conclusion that the effect of concomitant physical symptoms is both statistically and clinically meaningful. There is reason to assume that this effect might be substantial. One of the studies165 reported that for patients with a physical presentation of MDD at baseline the decrease on a scale measuring the severity of MDD was almost twice as much as for patients with a psychological presentation. Two studies163;164 reported odds ratios for an unfavourable course for patients with physical symptoms at baseline. The first163 reported an OR of 2.29 (CI 1.13-5.0). The second study164 used another scale to measure physical symptoms, the somatic scale of the SCL-90-R. An odds ratio over a comparable difference in points on this scale corresponds with an OR of 1.63 (CI 1.27-2.10). The odds ratio thus appears to lie between 1.27 and 5.0, implying that in the studies reporting odds ratios163;164 the likelihood of an unfavourable course of MDD was somewhere between 1.27 and 5.0 times as high for patients suffering from concomitant physical symptoms than for patients not suffering from such symptoms. This is a rough estimate, but it does show that the effect of physical symptoms on the prognosis of MDD might be quite large. This suggests that difficulty in translating efficacy findings to real-world settings is associated with increased suffering from concomitant physical symptoms in patients with MDD.

Limitations and strengths of the study

Because of the importance of the subject under study, it is somewhat puzzling that we could only find nine studies that met our inclusion criteria. The fact that unpublished studies were not available for consideration might possibly play a role. Studies that did not find the effect we describe might not have been published. Therefore, the conclusion of our best-evidence synthesis, i.e. that there seems to be strong evidence for a negative association between physical symptoms and the prognosis of MDD, is only a first step. Evidently, more studies are needed in order to gain more insight into the phenomenon.

Future studies could focus on effect-modification by physical symptoms of specific types of treatment for MDD. In particular, trials evaluating the effect of physical symptoms on psychotherapy could provide more insight. Studies which include a clinical
assessment as to whether or not the physical symptoms that depressed patients experience are medically unexplained would also be welcome. From the studies included in this review we were unable to make such a distinction. An overview of scales to assess physical symptoms, that could be included in baseline questionnaires, is provided by Hiller and Janca.171

Conclusions and implication for daily practice

Based on our results, a few recommendations can be made for daily clinical practice. Given the evidence for quite strong associations between physical symptoms and subsequent prognosis of depression, co-morbid physical symptoms should be assessed systematically in depressed patients. Since both adherence to pharmacotherapy168;169 and outcome seem to be negatively affected167, specific attention should be paid to physical symptoms -both explained and unexplained- in patients suffering from depression.

Effective treatment for MDD in the co-occurrence of physical symptoms may require collaboration between different health care professionals. The experiences in primary care with the collaborative care model are promising in this respect, since this method of organising co-operation between health care professionals has been found to be effective in the treatment of both MDD12 and persistent medically unexplained symptoms.14

Appendix: search strategy

<table>
<thead>
<tr>
<th>Category</th>
<th>Key words and medical subject headings within categories combined with the OR operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Depressive Disorder</td>
<td>Major Depression, Depression (Emotion), Depressive Disorder/therapy, Depression/therapy, Depressive Disorder/diagnosis, Depressive Disorder/drug therapy, Depressive Disorder/psychology</td>
</tr>
<tr>
<td>Information on the outcome or course of the disorder</td>
<td>Remission, relapse, outcome, prognosis, remit*, recur*, effectiv*, efficacy, recurrence, Treatment Outcome, residual symptom</td>
</tr>
<tr>
<td>Design</td>
<td>Clinical trial, trial, epidemiologic studies, epidemiologic, epidemiology, prognos*, prognosis, prediction</td>
</tr>
<tr>
<td>Physical symptoms</td>
<td>102 symptoms and symptom-related keywords were entered into this category, ranging from pain and fatigue to somatoform disorders, keywords such as somatisation or somatization, to symptoms such as headache, nausea and gas.</td>
</tr>
</tbody>
</table>

Combined with NOT category

Diabetes, Neoplasms, Cardiovascular Diseases, Lung Diseases, Obstructive Pulmonary Diseases
Acknowledgments

This project was funded by the ‘Innovatiefonds Zorgverzekeraars’ (Foundation for Innovation of the Health Care Insurances in the Netherlands).

The authors gratefully acknowledge the help received from Ingrid Riphagen in designing the search strategy.