APPENDIX A. CONDITION PRIORITIZATION INSTRUCTIONS

Please prioritize the list of health conditions by assigning 0 to 3 stars to items listed below. When prioritizing, consider the following: (1) prevalence of the condition in Veterans, (2) the burden of the illness, (3) priorities for gender-specific care for women Veterans (not Active Duty), and (4) conditions with known disparities between women Veterans and non-Veterans. Items are categorized by conditions affecting men and women compared with gender-specific conditions.

You are given a total of 11 stars to allocate to any of the 34 items listed in the 2 categories. You may use up to 3 stars per item. To add stars to a selection, position your mouse over the dots in the right hand column. More stars equal higher priority.

Alzheimer’s disease
Anxiety (ie, GAD, panic disorder)
Osteoporosis
Coronary artery disease (ie, chronic angina)
Coronary artery disease (ie, acute coronary syndrome/myocardial infarction)
Chronic obstructive pulmonary disease
Cerebrovascular disease (ie, ischemic stroke)
Depression (ie, MDD & dysthymia)
Diabetes mellitus-type2
Eating disorders
Connective tissue disease (ie, fibromyalgia)
Headache (ie, migraine)
Hepatitis-C
HIV
Hyperlipidemia
Hypertension
Irritable bowel syndrome
Incontinence
Insomnia
Joint disorders (ie, osteoarthritis: hip and knee)
Joint disorders (ie, rheumatoid arthritis)
Obesity/overweight
Chronic pain
Posttraumatic stress disorder
Spine disorders (ie, chronic low back pain)
Substance use disorder
Traumatic brain injury
Thyroid disorders
Tobacco use disorder

SEPARATOR BETWEEN SEX-SPECIFIC ITEMS
Contraceptive care
Infertility
Menstrual disorders (ie, abnormal uterine bleeding)
Menopausal disorders
Depressive disorders (ie, postpartum depression)
Depressive disorders (ie, premenstrual dysphoric disorder)
Prenatal care
APPENDIX B. SEARCH STRATEGIES

CONDITION: MAJOR DEPRESSIVE DISORDER

PubMed: Searched October 31, 2014

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>Search &quot;antidepressive agents&quot;[Pharmacological Action] OR &quot;antidepressive agents&quot;[MeSH Terms] OR &quot;antidepressive&quot;[tiab] OR antidepressant*[tiab]</td>
<td>141775</td>
</tr>
<tr>
<td>#5</td>
<td>Search #1 AND (#2 OR #3 OR #4 OR #5)</td>
<td>32393</td>
</tr>
<tr>
<td>#6</td>
<td>Search #5 AND (systematic[sb] OR &quot;Systematic Review&quot;[tiab] OR meta-analysis[tiab] OR &quot;meta analysis&quot;[tiab])</td>
<td>1792</td>
</tr>
<tr>
<td>#7</td>
<td>Search #6 NOT (&quot;Adolescent&quot;[Mesh] OR &quot;Child&quot;[Mesh] OR &quot;Adolescent&quot;[Mesh] OR &quot;Adult&quot;[Mesh])</td>
<td>1677</td>
</tr>
<tr>
<td>#8</td>
<td>Search #7 NOT (&quot;Animals&quot;[Mesh] NOT &quot;Humans&quot;[Mesh])</td>
<td>1677</td>
</tr>
<tr>
<td>#9</td>
<td>Search #7 NOT (&quot;Animals&quot;[Mesh] NOT &quot;Humans&quot;[Mesh]) Filters: published in the last 5 years</td>
<td>631</td>
</tr>
<tr>
<td>#10</td>
<td>Search #7 NOT (&quot;Animals&quot;[Mesh] NOT &quot;Humans&quot;[Mesh]) Sort by: Author Filters: published in the last 5 years; English</td>
<td>581</td>
</tr>
</tbody>
</table>

Cochrane Library: Searched October 31, 2014

<table>
<thead>
<tr>
<th>ID</th>
<th>Search</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>major depressive disorder:ti,ab,kw (Word variations have been searched)</td>
<td>2783</td>
</tr>
<tr>
<td>ID</td>
<td>Search</td>
<td>Results</td>
</tr>
<tr>
<td>-----</td>
<td>------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>#2</td>
<td>major depression:ti,ab,kw (Word variations have been searched)</td>
<td>3691</td>
</tr>
<tr>
<td>#3</td>
<td>major depression disorder:ti,ab,kw (Word variations have been searched)</td>
<td>24</td>
</tr>
<tr>
<td>#4</td>
<td>dysthmic disorder:ti,ab,kw (Word variations have been searched)</td>
<td>241</td>
</tr>
<tr>
<td>#5</td>
<td>dysthymia:ti,ab,kw (Word variations have been searched)</td>
<td>379</td>
</tr>
<tr>
<td>#6</td>
<td>involutional depression:ti,ab,kw (Word variations have been searched)</td>
<td>12</td>
</tr>
<tr>
<td>#7</td>
<td>involutional melancholia:ti,ab,kw (Word variations have been searched)</td>
<td>0</td>
</tr>
<tr>
<td>#8</td>
<td>involutional psychosis:ti,ab,kw (Word variations have been searched)</td>
<td>0</td>
</tr>
<tr>
<td>#9</td>
<td>involutional psychoses:ti,ab,kw (Word variations have been searched)</td>
<td>0</td>
</tr>
<tr>
<td>#10</td>
<td>(or #1-#9)</td>
<td>6077</td>
</tr>
<tr>
<td>#11</td>
<td>#10 Publication Year from 2009 to 2014, in Cochrane Reviews (Reviews only) and Other Reviews</td>
<td>117</td>
</tr>
</tbody>
</table>

### CONDITION: TYPE 2 DIABETES

**PubMed: Searched February 6, 2015**

<table>
<thead>
<tr>
<th>Search</th>
<th>Items found</th>
</tr>
</thead>
</table>
### Search Results

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>Search #1 AND #2</td>
<td>31000</td>
</tr>
<tr>
<td>#5</td>
<td>Search #1 AND #4</td>
<td>18805</td>
</tr>
<tr>
<td>#7</td>
<td>Search #1 AND #6</td>
<td>6871</td>
</tr>
<tr>
<td>#8</td>
<td>Search &quot;Weight Reduction Programs&quot;[Mesh] OR &quot;Weight Reduction Program&quot;[tiab] OR &quot;Weight control Program&quot;[tiab] OR &quot;Nutrition Therapy&quot;[Mesh] OR &quot;weight management&quot;[tiab]</td>
<td>83630</td>
</tr>
<tr>
<td>#9</td>
<td>Search #1 AND #8</td>
<td>2694</td>
</tr>
<tr>
<td>#10</td>
<td>Search &quot;Bariatric Surgery&quot;[Mesh] OR &quot;Bariatric Surgery&quot;[tiab]</td>
<td>17552</td>
</tr>
<tr>
<td>#11</td>
<td>Search #1 AND #10</td>
<td>1329</td>
</tr>
<tr>
<td>#13</td>
<td>Search #1 AND #12</td>
<td>3965</td>
</tr>
<tr>
<td>#14</td>
<td>Search #3 OR #5 OR #7 OR #9 OR #11 OR #13</td>
<td>41293</td>
</tr>
<tr>
<td>#15</td>
<td>Search #14 NOT (&quot;Adolescent&quot;[Mesh] OR &quot;Child&quot;[Mesh] OR &quot;Infant&quot;[Mesh]) NOT &quot;Adult&quot;[Mesh])</td>
<td>40423</td>
</tr>
<tr>
<td>#16</td>
<td>Search #15 NOT (&quot;Animals&quot;[Mesh] NOT &quot;Humans&quot;[Mesh])</td>
<td>37107</td>
</tr>
</tbody>
</table>

### CONDITION: CHRONIC PAIN

PubMed: Searched February 27, 2015

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>Search #2 NOT (&quot;Animals&quot;[Mesh] NOT &quot;Humans&quot;[Mesh])</td>
<td>2025</td>
</tr>
<tr>
<td>#4</td>
<td>Search #3 NOT (&quot;Adolescent&quot;[Mesh] OR &quot;Child&quot;[Mesh] OR &quot;Infant&quot;[Mesh]) NOT &quot;Adult&quot;[Mesh])</td>
<td>1985</td>
</tr>
<tr>
<td>#5</td>
<td>Search (&quot;2009/10/01&quot;[Date - Publication] : &quot;3000&quot;[Date - Publication]) AND #4</td>
<td>1145</td>
</tr>
</tbody>
</table>
Mapping the Evidence: Sex Effects in High-Impact Evidence-based Synthesis Program Conditions for Women Veterans

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR &quot;cognitive&quot;[tiab] OR &quot;behavior&quot;[tiab] OR &quot;behavioral&quot;[tiab] OR &quot;relaxation&quot;[tiab] OR &quot;acceptance&quot;[tiab])</td>
<td>196</td>
<td></td>
</tr>
<tr>
<td>#9 Search #5 AND (&quot;Anti-Inflammatory Agents, Non-Steroidal&quot;[Pharmacological Action] OR &quot;Anti-Inflammatory Agents, Non-Steroidal&quot;[Mesh] OR &quot;NSAIDs&quot;[tiab] OR &quot;Nonsteroidal Anti Inflammatory Agents&quot;[tiab] OR &quot;Nonsteroidal Anti Inflammatory Agent&quot;[tiab])</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>#10 Search #5 AND (&quot;Capsaicin&quot;[Mesh] OR &quot;Capsaicin&quot;[tiab])</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>#11 Search #5 AND (&quot;Lidocaine&quot;[MeSH Terms] OR &quot;lidocaine&quot;[tiab]) AND (&quot;transdermal patch&quot;[MeSH Terms] OR &quot;transdermal&quot;[tiab] OR &quot;patch&quot;[tiab])</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>#12 Search #5 AND (&quot;Antidepressive Agents&quot;[Mesh] OR &quot;Antidepressive Agents&quot;[Pharmacological Action] OR &quot;Duloxetine&quot;[tiab] OR &quot;Venlafaxine&quot;[tiab])</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>#13 Search #5 AND (&quot;pregabalin&quot;[Supplementary Concept] OR &quot;pregabalin&quot;[tiab])</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>#14 Search #5 AND (&quot;gabapentin&quot;[Supplementary Concept] OR &quot;gabapentin&quot;[tiab])</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>#15 Search #5 AND (&quot;Hyaluronic Acid&quot;[Mesh] OR &quot;Hyaluronic Acid&quot;[tiab])</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>#16 Search #5 AND (&quot;Steroids&quot;[Mesh] OR &quot;steroid&quot;[tiab] OR &quot;steroids&quot;[tiab])</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>#19 Search #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18</td>
<td>538</td>
<td></td>
</tr>
</tbody>
</table>

Cochrane Library: Searched February 27, 2015

<table>
<thead>
<tr>
<th>ID</th>
<th>Search</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>&quot;chronic pain&quot;:ti,ab,kw (Word variations have been searched)</td>
<td>2689</td>
</tr>
<tr>
<td>#2</td>
<td>&quot;fibromyalgia&quot;:ti,ab,kw (Word variations have been searched)</td>
<td>1249</td>
</tr>
<tr>
<td>#3</td>
<td>#1 or #2 Publication Year from 2009 to 2015, in Cochrane Reviews (Reviews and Protocols)</td>
<td>88</td>
</tr>
</tbody>
</table>
## APPENDIX C. ELIGIBILITY CRITERIA FOR SYSTEMATIC REVIEWS

<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Populations</strong></td>
<td><strong>Depressive disorders</strong>: Adults with major depressive disorder, persistent depressive disorder (dysthymia), or subsyndromal depression/minor depression/depression-NOS; also included are reviews broadly addressing effects of treatment in patients with co-occurring chronic medical illness&lt;br&gt;&lt;br&gt;<strong>Diabetes</strong>: Adults with type 2 diabetes mellitus&lt;sup&gt;a&lt;/sup&gt;</td>
<td><strong>Depressive disorders</strong>: Reviews focused on bipolar disorder, grief, premenstrual dysphoric disorder, psychotic depression, depression subtypes (eg, atypical depression; melancholic depression); also excluded are reviews focused on subsets of depressed patients who have a specific comorbid medical condition (eg, diabetes, heart disease) or psychiatric illness (eg, alcohol misuse)&lt;br&gt;&lt;br&gt;<strong>Diabetes</strong>: None</td>
</tr>
<tr>
<td><strong>Interventions</strong></td>
<td><strong>Depressive disorders</strong>&lt;sup&gt;b&lt;/sup&gt;:&lt;br&gt;• Antidepressants (SSRI, SNRI, TCA)&lt;br&gt;• Therapy: CBT, CT, IPT, MBCT, PST, ST, psychodynamic, reminiscence therapy delivered in person, in groups, or by internet&lt;br&gt;• Supervised exercise&lt;br&gt;• Guided self-help based on principles of CBT&lt;br&gt;• QI strategies: collaborative care, co-located care, women-only clinic&lt;br&gt;• Alternative: dietary supplements (eg, fish oil; vitamin D), yoga, acupuncture, St. John’s wort, SAM-e&lt;br&gt;• Medications: atypical antipsychotics, ketamine, adjunctive medications used for augmentation (eg, psychostimulants, thyroid hormone, lithium) that have not been specified as eligible medications; reviews of single medications, rather than a drug class, unless review is an individual patient data meta-analysis&lt;br&gt;• Somatic: electroconvulsive therapy, light therapy, transcranial magnetic stimulation, vagal nerve stimulation, deep brain stimulation&lt;br&gt;• Psychotherapies: dialectical behavioral therapy, music therapy, traditional long-term psychodynamic therapy, pet therapy&lt;br&gt;• Treatment sequencing (eg, switching antidepressants)&lt;br&gt;• Interventions to prevent depressive disorder (eg, interferon therapy for hepatitis C)&lt;br&gt;• Interventions to prevent diabetes&lt;br&gt;• Alternative: dietary supplements, acupuncture, meditation-based interventions (eg, transcendental meditation)</td>
<td></td>
</tr>
<tr>
<td>PICOTS Element</td>
<td>Inclusion Criteria</td>
<td>Exclusion Criteria</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
</tbody>
</table>
| shared medical appointments, chronic disease management (eg, telephone and internet-based interventions), stepped-care models, nurse-managed clinics, women-only clinic, smartphone applications | Chronic pain:  
- Antidepressants: SNRIs (duloxetine, venlafaxine, milnacipran), TCA, SSRI  
- Calcium channel 2 delta ligands*: pregabalin, gabapentin  
- Muscle relaxants*  
- Topical treatments: NSAIDs, capsaicin, lidocaine patch  
- Joint injection: steroid (if back, must be lumbar region)\(^\text{f}\), hyaluronic acid\(^\text{g}\)  
- Behavioral treatments focused on pain management: psychoeducation, CBT, mindfulness-based and acceptance-based therapy, relaxation therapy, biofeedback in groups or by internet  
- Exercise: aerobic, strengthening, or stretching performed with supervision (eg, physical therapist, pool therapy), as part of a class (eg, yoga class, tai chi), or as medically directed self-care  
- Integrative and complementary medicine*: acupuncture; spinal manipulation (chiropractic care)  
- QI strategies: multidisciplinary care (eg, multidisciplinary pain clinic), co-located care (eg, behaviorist in primary care), women-only clinic; telephone-based care  
- Self-management strategies used to decrease pain symptoms | Medications: any medication or class of medication not listed in the included section, insulin pump regimens, types or intensity of insulin regimens, colesteyl, alpha-glucosidase inhibitors, bromocriptine, miglitol; reviews of single medications unless the medication is insulin, metformin, repaglinide, nateglinide, pioglitazone, rosiglitazone, or pramlintide, or review is an individual patient data meta-analysis  
- Somatic: type or intensity of glucose monitoring  
- Surgical interventions other than bariatric surgery  
- QI: endocrinology clinics, QI interventions with clinician as intervention target (eg, decision support via computer reminders) |

Comparators | All conditions: Any active or inactive comparators | None |

---

* denotes conditions that have shown promise in preclinical or early clinical trials.
<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
</table>
| **Outcomes**   | All conditions: Patient health outcomes and adverse effects  
  • **Depressive disorders**: Depressive symptoms, functional status  
  • **Diabetes**: Glycemic control, weight, mortality, microvascular and macrovascular events, adverse effects  
  • **Chronic pain**: Pain severity, functional status | **Depressive disorders**: Provider outcomes, acceptance of intervention, prevalence of intervention, cost of intervention without reporting patient health outcomes  
  **Diabetes**: Provider outcomes, adherence to and/or acceptance of intervention, blood pressure, lipids, prevalence of intervention, cost of intervention  
  **Chronic pain**: Provider outcomes, adherence to and/or acceptance of intervention, prevalence of intervention; cost of intervention without reporting patient health outcomes |
| **Timing**     | Any duration of follow-up | None |
| **Settings**   | **Depressive disorders**: Any setting  
  **Diabetes**: Outpatient settings  
  **Chronic pain**: Outpatient settings | **Depressive disorders**: None  
  **Diabetes**: In-hospital setting, or focused on effects in low-income countries  
  **Chronic pain**: Focused on work-based programs |
| **Design**     | All conditions: Systematic reviews or individual patient level meta-analyses; must have search strategy, eligibility criteria, and analysis/synthesis plan  
  • **Depressive disorders**: May include RCTs (antidepressants, therapy, exercise), or quasi-experimental studies (QI interventions), or observational studies (if focused on adverse effects)  
  • **Diabetes**: May include RCTs (medications, behavioral interventions), or quasi-experimental studies (QI interventions), or observational studies (if focused on adverse effects)  
  • **Chronic pain**: May include RCTs (medications, behavioral, exercise), or quasi-experimental studies (QI strategies), or observational studies (if focused on adverse effects) | Nonsystematic reviews |
<p>| <strong>Publications</strong> | English-language only | Non-English publications |</p>
<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Published October 2009 to present</td>
<td>Publications before October 2009</td>
</tr>
</tbody>
</table>

a Mixed diabetes populations were included if patients with type 2 diabetes were analyzed separately.

b Interventions may be for acute-phase treatment, treatment-resistant depression, or maintenance.

c Includes tai chi, Pilates, yoga, and related forms of exercise.

d Includes supervised programs that use changes in physical activity, diet, or a combination of these approaches to achieve weight change or improved glycemic control.

e For chronic low back pain and fibromyalgia only.

f For chronic low back pain and chronic knee pain only.

g For chronic knee pain only.

h Includes stroke, cardiac event (eg, myocardial infarction), nephropathy, neuropathy (including diabetic foot ulcer), and changes in cognition.

i Includes cancer, osteoporosis, hypoglycemia, changes in cognition, lactic acidosis, adverse gastrointestinal effects, and cardiovascular and serious adverse events.

Abbreviations: CBT=cognitive behavioral therapy; CT=cognitive therapy; DPP-4=dipeptidyl peptidase 4; GLP-1=glucagon-like peptide-1; IPT=interpersonal therapy; MBCT=mindfulness-based cognitive therapy; NOS=not otherwise specified; NSAID=nonsteroidal anti-inflammatory drug; PICOTS=population, intervention, comparator, outcome, timing, setting; PST=problem-solving therapy; QI=quality improvement; RCT=randomized controlled trial; SAM-e=S-adenosylmethionine; SNRI=serotonin-norepinephrine reuptake inhibitor; SSRI=selective serotonin reuptake inhibitor; ST psychodynamic=short-term psychodynamic; TCA=tricyclic antidepressant; TENS=transcutaneous electrical nerve stimulation
## APPENDIX D. RESPONSES TO REVIEWER COMMENTS

### Mapping the Evidence: Sex Effects in High-impact Conditions for Women Veterans – Depression, Diabetes, and Chronic Pain

<table>
<thead>
<tr>
<th>Question Text</th>
<th>Reviewer Number</th>
<th>Comment</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are the objectives, scope, and methods for this review clearly described?</td>
<td>1</td>
<td>Yes</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Yes</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Yes</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Yes</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Yes</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Yes</td>
<td>Acknowledged</td>
</tr>
<tr>
<td>Is there any indication of bias in our synthesis of the evidence?</td>
<td>1</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td>Are there any published or unpublished studies that we may have overlooked?</td>
<td>1</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>No</td>
<td>Acknowledged</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Yes - It is difficult to know whether studies were overlooked, give the way results are reported in broad categories, but the evidence reviewed does not appear to be comprehensive or include the most high impact/high quality reviews available. For example, only one review of medications for low back pain is discussed (on page 22, citation 193). Other reviews of medications for back pain are available and seemingly should be included, but I can't tell if they were because the number of reviews of medications is presented for all chronic pain conditions combined. For LBP injections, 2 reviews from low tier journals are highlighted, but the high quality American Pain Society sponsored reviews of interventional therapies by Roger Chou et al (published in Journal of Pain in 2009) are not included.</td>
<td>We identified 3 Chou citations in the <em>Journal of Pain</em> during 2009. These were examined and found not eligible. However, we identified another review from this author in <em>Spine</em> (2009), which has been included in the final report. We also conducted a supplemental search for studies that addressed eligible interventions for chronic knee pain due to osteoarthritis. Additional eligible reviews were identified and are included in the final report. We think the relatively low number of eligible articles for chronic pain conditions is in part due to the interventions chosen for evaluation.</td>
</tr>
<tr>
<td>Question Text</td>
<td>Reviewer Number</td>
<td>Comment</td>
<td>Response</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>-------------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>Methodology seems very sound - further examination of primary trials is valuable and a good plan.</td>
<td>Acknowledged</td>
</tr>
<tr>
<td>7</td>
<td>No</td>
<td></td>
<td>Acknowledged</td>
</tr>
<tr>
<td>8</td>
<td>No</td>
<td></td>
<td>Acknowledged</td>
</tr>
<tr>
<td>Additional suggestions or comments can be provided below. If applicable, please indicate the page and line numbers from the draft</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Methodology seems very sound - further examination of primary trials is valuable and a good plan.</td>
<td>Thank you.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Overall feedback: This review seeks to create “...an evidence map to aid prioritization and development of implementation projects and research initiatives” in order to address the emerging healthcare needs of women Veterans, which has the potential to be of great benefit to policy makers and researchers alike. Overall the manuscript flows exceedingly well and the authors present the information in a balanced, logical, thoughtful, and succinct manner. Particular strengths are the inclusiveness of the treatments types that were reviewed and the consistent organization and framing that facilitates readership from section to section. Overall this reviewer found that the manuscript may be enhanced by some minor adjustments or additions in the following specific areas: 1. The title, “Women’s Health Evidence Map”, is rather broadly worded given that the paper focuses on three health conditions (depression, DM, and chronic pain). This reviewer recommends re-working the title or adding a subtitle that speaks to the content area covered in the report. A subtitle may be particularly suitable if subsequent papers may expand the health evidence map by reviewing additional health conditions. 2. In various locations the studies contained within are referred to descriptively as “low grade”, “low quality”, “high quality”, “good quality” and so forth. There is a paragraph within the paper that indicates that a formal</td>
<td>Thank you.</td>
<td></td>
</tr>
</tbody>
</table>

We agree that the title was incomplete. It has been revised to “Mapping the Evidence: Sex Effects in High-impact Conditions for Women Veterans—Depression, Diabetes, and Chronic Pain.”

We’ve standardized the descriptors and clarified when we are referring to a high-quality review (as defined by a Cochrane, ESP, or EPC review) versus a summary statement about the quality of the evidence made by the
<table>
<thead>
<tr>
<th>Question Text</th>
<th>Reviewer Number</th>
<th>Comment</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>assessment of methodological rigor is “beyond the scope of the project” (see Quality Assessment page 5) however subsequent use of these descriptors suggests at least informal evaluation of the papers. This reviewer suggests including a description (perhaps in the Quality Assessment section) that at least informally anchors these descriptors. In addition, would recommend standardizing the descriptors—e.g., using low quality or low grade unless these two descriptors differ from one another.</td>
<td></td>
<td>review authors.</td>
<td>We’ve added additional detail to the methods to describe the stakeholders and the process for prioritizing the included conditions.</td>
</tr>
<tr>
<td>3. The description of the iterative process and use of stakeholders to determine the health conditions chosen left this reviewer looking for additional information. For example, what was the composition of the stakeholder group? Why only three conditions versus some other number? Did the iterative process yield agreement/ranking that only suggested review of 3 health conditions with a fourth (or fifth, etc..) not reaching a reasonable consensus or convergence of opinion?</td>
<td></td>
<td></td>
<td>Edit made.</td>
</tr>
<tr>
<td>4. Diabetes Section: page 16 the follow sentence may be improved by revising “last” to “lastly”: “Last, all 3 reviews of DPP-4 inhibitors applied IPD meta-analyses to data from industry studies.137,142,143.”</td>
<td></td>
<td></td>
<td>We have moved this paragraph to the recommended section.</td>
</tr>
</tbody>
</table>
| 5. Evidence Gaps for Sex Effects Subsection: The paragraph that starts “In response to underrepresentation of women and minorities...” while containing important information, seems out of place in a summary of the evidence gaps for sex effects and more suited to the subsequent section entitled “Achieving Adequate Representation of Women in Clinical Studies”. The paragraph on “evidence gaps” may be enhanced by further description of where sex effects lead to reasonably sound information (information of use clinically) or to highlight what guidance is lacking with regard to clinical care given the very limited information regarding sex effects. | | | We considered the suggestion to add recommendations for clinical care. First, our goal was more descriptive. We also did not conduct a quality assessment of the systematic reviews. Furthermore, the sex results were generally small and arrived at by poorly suited analytic techniques. Thus, we do not feel confident in making specific clinical treatment recommendations. However, in order to address the clinical implications of our results, we have identified promising areas that could, with further research, lead to sex-specific clinical recommendations. In addition, we offer some discussion of how these proposed
6. Achieving Adequate Representation of Women in Clinical Studies Subsection: It seems implicit in this paragraph that inclusion of women Veterans in research studies is somehow only available for women Veterans who are enrolled in VA care (i.e., the sentence on page 27 reads: “Inclusion of women in VA research is particularly challenging because the enrollment of women Veterans in VHA, while growing, remains a relatively low proportion (6.5% in 2012).”). One could argue that adequate representation is also hampered by only enrolling women in VA care (into studies) rather than expanding recruitment to women Veterans more broadly. In addition, adequate representation of the needs of women Veterans more broadly is also affected by only studying only/primarily women enrolled in VA care. This portion of the paper may be improved by acknowledgement of these limitations and how these limitations may affect or influence the knowledge base.

7. Prioritizing Areas for Evaluation of Sex Effects Subsection: most of this paragraph appears to speak to the prior sub-section content, which is Achieving Adequate Representation of Women in Clinical Studies. This reviewer suggests relocating this information to the above section and adding additional text in the Prioritizing Areas for Evaluation of Sex Effects that that includes suggestions that speak to the three content areas (depression, diabetes, chronic pain) that were so thoroughly reviewed in earlier portions of the paper.

8. The paper would benefit from additional discussion points that address or highlight:
   a. Key points for care based on what is known now—as limited as that may be (perhaps in the Discussion or even more suited for the Evidence Gaps subsection on page 26).

We addressed this suggestion by broadening our discussion regarding barriers to recruitment, so that we noted (1) the small proportion of total veteran population who are women and (2) general barriers found for study participation by minorities and women. Thus, while we agree that expanding to women Veterans not enrolled in VA may increase the participant pool somewhat, we think this would ultimately not make a substantial difference. Additionally, recruiting these women poses significant complications for VA-based research, given that such projects often utilize VHA resources (e.g., EHR).

We appreciate this suggestion but have intentionally avoided statements on clinical implications. The primary goal was to describe the volume of studies and whether sex effects have been examined. Without a careful quality assessment, we are hesitant to suggest clinical actions.
<table>
<thead>
<tr>
<th>Question Text</th>
<th>Reviewer Number</th>
<th>Comment</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>as a result of the health evidence map. Suggesting content along the lines of clinical implications.</td>
<td></td>
<td>Based on the estimates of intervention effects that we report.</td>
<td></td>
</tr>
<tr>
<td>b. Future directions/research gaps. This could be a separate section or become part of a slightly re-worked section that is currently titled: Prioritizing Areas for Evaluation of Sex Effects. The content of that “prioritizing areas” is broad and speaks to the pressing need for researchers to incorporate adequate methodology to address sex effect questions within their projects. The content I would recommend adding (here or elsewhere) are gaps within the content areas of depression, DM, and chronic pain that would provide recommendations for the scientific community on possible future directions.</td>
<td></td>
<td>Thank you for this suggestion. We’ve added example of gaps and types of studies that could be conducted to address these gaps for the conditions reviewed. We’ve also discussed the potential clinical implications of results from these proposed studies.</td>
<td></td>
</tr>
<tr>
<td>In general, I found this evidence map to be difficult to follow and to review. I believe this is attributable primarily to the enormous scope of the review. Attempting to provide brief overviews of the evidence for all of the interventions in all three huge clinical areas has resulted in extremely superficial and potentially misleading summaries of evidence for the individual interventions and conditions. The overview approach also obscures and detracts from the findings on sex effects, which is the main focus and strength of the report. (For example, the evidence for medications in chronic LBP, see specific comment below.) I think these problems could be substantially addressed by focusing just on evidence (and lack thereof) regarding sex differences. Specific comments: 1. The definition and merits of individual patient data (IPD) meta-analysis should be briefly described (and citations provided) in the methods section. Also, as IPD is an unfamiliar abbreviation, it should be included in the abbreviation list and spelled out in the table footnotes.</td>
<td></td>
<td>Thank you for this comment. We agree that the scope is large and the key messages could be difficult to identify with the current report structure. Therefore, we’ve substantially reorganized the report to consolidate the findings on sex effects across the three conditions. We’ve moved the information on intervention effects in general to an Appendices (by condition), as these were not the primary focus of our evidence map. We’ve added a definition for IPD meta-analysis to Table 2 (Definitions of statistical approaches) and added a brief rationale for the merits of this approach to the methods section. We’ve been careful to define this abbreviation when used in tables.</td>
<td></td>
</tr>
<tr>
<td>Question Text</td>
<td>Reviewer Number</td>
<td>Comment</td>
<td>Response</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>2. The choice to limit reviews to the past 5 years may create an unintended bias for lower quality reviews in fields with few large funded RCTs (such as chronic low back pain). If a definitive high quality review has been published and no new trials are available, subsequent reviews are unlikely to be higher quality.</td>
<td></td>
<td>Limiting reviews to the past 5 years was based on the rationale that reviews older than this are often out of date. Although we agree this decision could lead to missing reviews on some interventions (eg, no recent published review), we disagree that it would introduce a bias toward selecting low-quality reviews.</td>
<td></td>
</tr>
<tr>
<td>3. Brief presentation of review findings without any details or quality assessment is potentially misleading. In the pain section in particular, included reviews have limitations that are not acknowledged and some of the best quality reviews seem to be absent.</td>
<td></td>
<td>We agree that brief presentations of intervention effects without quality assessment could be misleading. We’ve addressed this in several ways. First, we reorganized the report to emphasize sex effects, where we offer more information on review author’s comments on the quality of the evidence, and provide discussion of the limits of certain analytic techniques (ie, meta-regression vs IPD meta-analyses). We have also added statements in the Methods and Results sections cautioning readers that the reviews have not been assessed for quality and thus estimates of effect may be wrong.</td>
<td></td>
</tr>
<tr>
<td>4. Page 22, line 41-42: The two reviews were not of “joint injections.” One reviewed facet joint injections and the other reviewed epidural injections. Also, I believe the authors meant to say “reduction in pain” or “improvement in pain relief”, rather than “reduction in pain relief.” Without information about the quality of the reviews, the comparisons, the specific pain indication, or the time frame of the outcomes, the summary statement about improvement may be misleading. (In general, very short term improvement in pain is not considered clinically important in the setting of chronic pain.)</td>
<td></td>
<td>We had defined the term “joint injections” loosely and have now specified these injections as facet and epidural. We have made the edit to “reduction in pain.”</td>
<td></td>
</tr>
<tr>
<td>5. Page 22, lines 54-59: The description of this meta-analysis is confusing. It was an indirect (no direct comparisons) meta-analysis of 15 placebo-controlled trials that was sponsored by Eli Lilly with the purpose of comparing duloxetine to other analgesics. It included trials of several typical opioids (“narcotics” “scheduled opioids”), tramadol (“non-scheduled opioids”), duloxetine, paroxetine</td>
<td></td>
<td>We agree that without information about quality, etc., summary statements may be misleading. Therefore, throughout the report, information on general (ie, non-sex-specific) effects of interventions has been moved to Appendices, partially due to several reviewers’ correct observation that it had not been collected as rigorously.</td>
<td></td>
</tr>
</tbody>
</table>

This information has been moved to an appendix and therefore gets less emphasis. We added the modifier that this was an industry-sponsored study. We did not add an extensive discussion of the strengths and weaknesses of the review, since we did not quality-rate the studies or include this type of discussion for other reviews.
<table>
<thead>
<tr>
<th>Question</th>
<th>Reviewer</th>
<th>Comment</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>Number</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“SSRIs”), and etoricoxib (“NSAIDS”). For a variety of reasons (controversial methods, differences in trial designs between drugs), their findings require additional discussion and context. Also, this should not be the only systematic review of medications for LBP.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>page 7: I had some trouble following the literature flow chart (Figure 1). I wonder if it would be better to break down the figure into 3 smaller figures for depression, diabetes and chronic pain. Page 10: did any of the depression studies report on menopausal status? page 25: the second paragraph notes the percentage of studies evaluating sex as a moderator (19%; 9%; 5%). I think this information should be added to the abstract.</td>
<td>We considered this suggestion but decided a single figure was preferable. However, we revised the figure for clarity. This is an interesting question but we did not abstract data on menopausal status. We have added this information to the Abstract.</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Thank you for this very thorough review. I can only say that I am disappointed in the findings (that very few studies addressed gender differences) not in the quality of the report which was very complete. They only thing I might have considered is including a longer time frame since the last 5 years seemed very short, especially for pain and depression related findings.</td>
<td>We agree it is disappointing that so few studies address sex effects. Limiting the review to the past 5 years was based on the rationale that reviews older than this are often out of date. At this time, we are not able to extend the timeframe for included systematic reviews.</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>This report is very clearly written and laid out and tackles an important issue within and outside the VA, which is the extent to which sex/gender differences are examined systematically as well as the potential for the larger non-VA literature to help identify opportunities for leveraging national investments in research for use in improving women Veterans’ care in the VA health care setting. The review also focuses on top priority conditions of high prevalence and quality of life impact for men and women, offering opportunities to consider differences and potential disparities in a broad way even though this is focused and</td>
<td>Thank you for these comments.</td>
</tr>
</tbody>
</table>
### Question Text

framed as a Women's Health Evidence Map. Given women’s numerical minority in VA, this framing is appropriate.

The writing and methods are clear and easy to read and follow. The material provided is solid and usefully displayed. Modest issues arise in use of abbreviations throughout the tables and figures, which warrant attention. For example, Figures 4 and 5 on pages 14 and 15 have "psyched" as an abbreviation and while this is likely for psychoeducation, the "psyched" abbreviation has other meanings in our vernacular so some other abbreviation without alternative interpretation would be helpful.

The Discussion (and thus Abstract in part) would benefit from clearer presentation of the main takeaway messages. Emphasis appears, perhaps unintentionally, to be on studies often showing greater benefit in women. However, the first takeaway point would seem to take precedence, *ie*, that a remarkably small fraction of the investment in these trials has been on examining sex/gender differences. The policy implications, however, of readers focusing on the positive findings in that small fraction that examined these issues are potentially very significant and may suffer from publication bias and other issues. In other words, the efforts to require inclusion of women in trial enrollment, analysis and reporting has been underway for decades, with substantial evidence that women's response to pharmaceutical agents, for example, can be meaningfully clinically different than men's. So your primary takeaway message appears to be that this emphasis has failed despite regulatory requirements. Very little time is spent on this failure in the Discussion so the findings lack contextualization. Instead, much more time is spent on the summary of sex effects in the very small fraction of studies that bothered to report on sex/gender differences, and given publication biases toward positive results, we lack adequate knowledge of what might have been found in the

### Response

We were able to spell out terms within the figures, and included abbreviation callouts where needed.

We’ve revised the discussion to address the issues you raise. These changes include moving the summary of sex effects to the results section, moving the table of gaps in evidence to an earlier section of the discussion, and giving more emphasis in the text to the general paucity of studies evaluating sex-specific effects.
large majority that failed to examine these issues. The findings that were gleaned from the review are nonetheless of value, but they are not the whole story—the concern is that they will become the whole story for those looking to avoid the methodological investment of time and resources needed to more definitively characterize intervention effects by subgroup.

Page 27: Very little information is included with respect to what VA has done to increase inclusion of women Veterans in VA research through the Women's Health Research Network, which includes consortium development (i.e., training of VA investigators on issues around women Veterans' health and health care; training in oversampling techniques and subgroup analysis approaches) and development of a national practice based research network that facilitates inclusion of women by developing local VA facility capacity to identify, recruit and retain women in VA research. This network is probably worth mentioning/citing, as it is also now responsible for increasing inclusion on women in the VA Cooperative Studies Program. Suggest you consider the following references to support building up this particular and brief paragraph:


Thank you for your comment and suggestions. We have expanded our discussion of VA initiatives to encourage research on women’s health topics and improve inclusion of women Veterans in VA research.
<table>
<thead>
<tr>
<th>Question Text</th>
<th>Reviewer Number</th>
<th>Comment</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yano EM. A partnered research initiative to accelerate implementation of comprehensive care for women Veterans: The VA Women’s Health CREATE. Med Care. 2015 Apr;53(4 Suppl 1):S10-14.</td>
<td></td>
<td>I think this report is going to have substantial impacts within and outside the VA both for advancing the tenets of equitable benefits of our research investment by gender and for informing ongoing opportunities for reducing gender disparities and also for targeting scarce resources more efficiently when differences do not indeed exist.</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX E. OVERALL EFFECTS OF SELECTED INTERVENTIONS: DEPRESSION

The information provided in this appendix reflects what was reported by the authors of the original, primary publication. We have not separately assessed the article for quality or accuracy.

Effects by Diagnostic Group for Depressive Disorders

**Major depressive disorder or depressive disorder.** Antidepressants were evaluated using IPD meta-analysis, qualitative, and using mixed approaches. A review of 165 antidepressant trials in patients with major depressive disorder found clinical response rates of 54.3% for antidepressants and 37.9% for placebo. The most comprehensive comparative effectiveness review (n=234 studies) used a multiple treatment comparison analysis to compare different antidepressants. This study, conducted by an EPC, found that antidepressants were effective and that there were no clinically important differences in efficacy or effectiveness between antidepressants for acute- or continuation-phase treatment. Subgroup analyses did not show differences in treatment effects by age, sex, ethnicity, or comorbid condition.

Five reviews evaluated antidepressants for late-life depression. Results were inconsistent with some reviews finding antidepressants less effective for older adults (>65 years of age) than for younger adults. Other reviews found no moderating effect of older age. Antidepressants were associated with decreased suicidal risk in older (≥65 years of age) and younger adults.

**Psychotherapy** for depressive disorders was examined in 35 reviews. Reviews evaluating multiple therapies (n=92-132 included trials) found psychotherapy compared with inactive control more effective for reducing depressive symptoms (n=117; SMD 0.67, CI 0.60 to 0.75), and psychotherapy increased the proportion of patients with a clinical response or clinical remission. Therapies examined included cognitive behavioral therapy (CBT), problem-solving therapy (PST), short-term psychodynamic therapy, interpersonal therapy, behavioral activation, and nondirective counseling. However, when accounting for publication bias (SMD 0.42) or when restricting analyses to the highest quality studies (Cohen’s d 0.22), effects were diminished. Effects did not differ significantly for different types of psychotherapy. Therapy was more effective than control for improving hopelessness but not suicidality. For major depressive disorder and subsyndromal depression, psychotherapy did not differ from antidepressant medication but was less effective than antidepressants in patients with dysthymia.

Four reviews evaluated psychotherapy in older adults, finding the strongest evidence of efficacy for CBT (n=14; Hedges’ g -0.57, CI -0.80 to -0.34). Reviews concluded that CBT is effective in individual and group formats.

Four reviews evaluated computerized therapy (predominately based on CBT) compared with inactive control or treatment as usual for patients with depressive disorders. Consistent with the other reviews, the largest identified 19 RCTs and found computerized therapy effective for depressive disorders (n=19; Cohen’s d -0.56, CI -0.71 to -0.41). An ESP review found that 6 to 8 sessions of CBT or PST improved depressive symptoms more than control (n=7; SMD -0.42, CI -0.74 to -0.10) for patients recruited from primary care settings. Cochrane reviews concluded that
low-quality evidence suggested that third-wave therapy (e.g., behavioral activation, acceptance and commitment therapy) may be effective and similar to other psychotherapies.\textsuperscript{32,35}

Seven reviews evaluated a combination of antidepressants plus psychotherapy compared with one of these treatments alone for acute-phase treatment of depressive disorders.\textsuperscript{55,234-239} The largest review found greater reduction in depressive symptoms for combined treatment than for antidepressants alone (n=23; Hedges’ g 0.43, CI 0.29 to 0.57).\textsuperscript{237} Another review found a small benefit from combined treatment compared with psychotherapy plus placebo pill (n=16; Cohen’s d 0.25, CI 0.03 to 0.46).\textsuperscript{235}

Seven reviews evaluated aerobic exercise compared with other physical activity, antidepressants, or treatment as usual.\textsuperscript{240-246} Exercise was more effective than control with effect sizes ranging from small to moderate.\textsuperscript{243-246} Aerobic exercise was not more effective than other physical activity (n=8; SMD 0.01, CI -0.23 to 0.24) or antidepressants.\textsuperscript{240} However, aerobic exercise improved depressive symptoms when used to augment treatment as usual (n=4; SMD -0.44, CI -0.79 to -0.09).\textsuperscript{240,241} A single review of yoga\textsuperscript{242} found evidence that yoga was more effective than usual care. These reviews did not examine sex effects.

A single IPD meta-analysis evaluated low-intensity self-help interventions for depressive disorders or mixed depression and anxiety.\textsuperscript{41} Self-help interventions were more effective than treatment as usual for reducing depressive symptoms (n=16 comparisons; SMD -0.42, CI -0.55 to -0.29). An analysis of potential moderators of intervention effects suggested that intervention effects did not vary by sex but were greater as baseline severity increased.

Three reviews evaluated quality improvement interventions that were predominately a form of collaborative care.\textsuperscript{40,59,247} The largest review\textsuperscript{247} identified 40 trials, finding moderate benefit on depressive symptoms for collaborative care compared with treatment as usual or other enhanced care (Cohen’s d 0.31, CI 0.16 to 0.47).

**Dysthymia/persistent depressive disorder.** Six reviews evaluated interventions for dysthymia.\textsuperscript{51,55,204,248-250} A multiple treatment comparison meta-analysis found antidepressants more effective than placebo (n=45; OR for response 2.35 to 6.98 for various antidepressants).\textsuperscript{249} A conventional meta-analysis also found antidepressants more effective than placebo (n=9; RR for response 1.75, CI 1.49 to 2.04).\textsuperscript{204} There were no differences in treatment response between SSRIs and tricyclic antidepressants, but dropouts were fewer with SSRIs (n=7, RR=0.41; CI 0.19, 0.86).\textsuperscript{250} Psychotherapy was more effective than control for reducing depressive symptoms (n=8; Cohen’s d 0.23, CI 0.06 to 0.41) but less effective than antidepressants (n=10, Cohen’s d -0.31, CI -0.53 to -0.09).\textsuperscript{55,248} In one review, interpersonal psychotherapy plus antidepressants was more effective than antidepressants alone (OR 1.83),\textsuperscript{249} but in another review, combined treatment was not more effective than antidepressants alone for patients with chronic major depressive disorder, dysthymia, double depression, or recurrent depression (n=8; RR for response 1.20, CI 0.98 to 1.48).\textsuperscript{51} No other interventions were evaluated.

**Subsyndromal/minor depressive disorder.** Two reviews evaluated psychotherapy in older adults with subsyndromal depression.\textsuperscript{251,252} In the larger review,\textsuperscript{252} psychotherapy, including CBT, PST, and behavioral activation, was more effective than wait-list, placebo pill, or treatment as usual controls (n=5; 701 patients; no effect estimate). A single review\textsuperscript{253} identified 8 trials evaluating antidepressants for subsyndromal depression or mild major depressive disorder. Antidepressants showed small improvements for depressive symptoms compared with placebo.
Mapping the Evidence: Sex Effects in High-Impact Evidence-based Synthesis Program Conditions for Women Veterans

(n=2; MD Hamilton Depression Rating Scale -1.39, CI -2.41 to -0.36). None of these reviews evaluated sex effects, and no other interventions were evaluated.

**Treatment-resistant depression.** Three reviews evaluated interventions for treatment-resistant depression and found the evidence sparse. Mindfulness-based cognitive therapy plus treatment as usual was more effective than treatment as usual in 2 trials (MD on Beck Depression Inventory -10.28, CI -17.18 to -3.38). Compared with active management, an ESP review identified 2 good-quality trials that showed similar benefit from augmenting antidepressant treatment with other antidepressants or psychotherapy, or switching antidepressants. None of these reviews evaluated sex effects, and no other interventions were evaluated.

**Relapse prevention.** Seven reviews evaluated interventions for treatment-resistant depression. Psychotherapy alone or in combination with an antidepressant decreased the risk of relapse (n=8; RR=0.79, CI 0.66 to 0.96). Specific therapies that were found effective were group or individual CBT, interpersonal therapy, and mindfulness-based cognitive therapy. Antidepressants (primarily SSRIs and SNRIs) were effective in preventing relapse (n=54; OR for relapse 0.35, CI 0.32 to 0.39) in patients who had responded to an antidepressant. An EPC-conducted multiple treatment comparison meta-analysis found no clinically important differences in efficacy between different antidepressant medications. None of these reviews evaluated sex effects, and no other interventions were evaluated.
APPENDIX F. OVERALL EFFECTS OF SELECTED INTERVENTIONS: TYPE 2 DIABETES

The information provided in this appendix reflects what was reported by the authors of the original, primary publication. We have not separately assessed the article for quality or accuracy.

Effects by Interventions for Diabetes

**Medications.** Of 120 eligible reviews on medications, we performed full data extraction on 75, prioritizing those that evaluated multiple drug classes or addressed less frequently examined categories (eg, meglitinides), among other criteria (see Methods). Glycemic control (n=38; reviews) and weight (n=25) were the most frequently included outcomes. Cardiovascular events (n=12) and mortality (n=16) were evaluated less often. Among adverse events, hypoglycemia was the most frequently discussed (n=29). The largest review, originating from an organization known for high-quality reviews, focused on risk for lactic acidosis associated with metformin therapy and included 347 studies.78 This study evaluated 70,490 patient-years of metformin use compared with 55,451 patient-years among nonusers, with 39% of women among both users and nonusers, and found no reported cases of lactic acidosis. The upper bound for the true incidence was estimated to be 4.3 cases per 100,000 patient-years among metformin users compared with 5.4 cases for nonusers.78

The next largest review included 277 RCTs, with 46% women overall, and examined effectiveness (dichotomously defined as reaching hemoglobin A1c target or not) and risk for adverse events for multiple drug classes compared with metformin.260 This review used network meta-analysis, conducting 173 direct and 180 indirect comparisons. Most drug classes were found to improve the likelihood of achieving glycemic targets when added to existing metformin therapy, ranging from the lowest benefit for rosiglitazone (OR 1.2, CI 1.1 to 1.3) to the highest for glucagon-like peptide 1 agonists (OR 11.1, CI 3.4 to 35.9). Sulfonylureas (OR 3.95, CI 1.82 to 8.55), α-glycosidase inhibitors (OR 3.24, CI 1.69 to 6.24), and meglitinides (OR 3.25, CI 1.83 to 5.75) were all associated with increased risk for hypoglycemia when added to metformin.260

**Bariatric surgery.** Of 12 reviews on bariatric surgery, only 3 were restricted to RCTs.261-263 Weight loss and diabetes remission were evaluated by nearly all reviews. The 3 largest and most recent reviews included 39 studies,100 35 studies,264 and 33 studies.265 One of these reviews examined multiple surgical procedures and reported sex effects on diabetes remission (key results reported above).100 The other reviews evaluated gastric banding versus multiple comparators,264 and gastric bypass versus sleeve gastrectomy.265 Gastric banding was associated with average 47% excess weight loss over 2 years.264 There were no differences in rates of diabetes remission (67% and 81% at 3 and 36 months for gastric bypass, compared with 56% and 80% for sleeve gastrectomy) or in excess weight loss between the 2 procedures.265

**Psychoeducation and mixed behavioral interventions.** Six reviews investigated the effect of psychoeducation and/or mixed behavioral interventions on glycemic control (n=6 reviews) and/or weight (n=3 reviews). Included studies ranged from 13 to 33, and 4 reviews were restricted to RCTs.79,83,106,266 Three reviews reported proportion of women included in primary studies (range 45% to 70%),102,266,267 and one specifically evaluated treatment effects in minority women.102
Four reviews evaluated mixed interventions consisting of psychoeducation and exercise with or without dietary advice.\textsuperscript{79,83,102,106} The other 2 mixed interventions included mobile phone applications\textsuperscript{267} and internet-based self-management strategies.\textsuperscript{266} The largest review (n=33 studies) was conducted by the Cochrane collaboration, and examined culturally appropriate health education and exercise in ethnic minority groups in middle- and upper-income countries.\textsuperscript{83} There was high-quality evidence to support that these interventions reduced HbA1c at 6 months (n=14; MD -0.53%, CI -0.72% to -0.35%), with some sustained effect at 12 months (-0.2%) and 24 months (-0.3%). There were no statistically significant effects on body mass index. Subgroup analyses were not performed. Another review, based on a 2011 AHRQ report, inspected the progression of diabetes-related complications and found significantly decreased risks of peripheral neuropathy, nonfatal myocardial infarction, and death with lifestyle interventions at 13.3 years of follow-up.\textsuperscript{79} There were no significant differences in risk for retinopathy or nephropathy between the intervention and control groups.

One review of internet-based self-management provided qualitative results showing that psychoeducation and behavioral techniques via mobile phone applications were associated with improved HbA1c (n=10; WMD -0.81, CI -1.11% to -0.50%).\textsuperscript{267}

**Supervised exercise.** Of 14 reviews on supervised exercise, 10 were restricted to RCTs, and the number of included studies ranged from 2 to 34. Six systematic reviews reported proportions of women in primary studies (median 53%, range 3% to 88% female), but none reported sex effects. Nearly all reviews (n=13) reported results for glycemic control, while 4 reported on weight loss,\textsuperscript{74,268-270} and 6 reported other outcomes.\textsuperscript{74,269,271-274} None reported adverse events.

Five reviews compared the effects of aerobic exercise, resistance exercise, and combinations of the 2 on HbA1c.\textsuperscript{268,269,275-277} The largest and most recent review restricted to RCTs included 26 primary studies with a minimum intervention duration of 12 weeks.\textsuperscript{276} This review found that all forms of exercise reduced HbA1c (aerobic MD -0.70%, CI -1.02% to -0.38%; resistance MD -0.62%, CI -1.14% to -0.11%; combined MD -0.47%, CI -0.64% to -0.31%) and that exercise volume (ie, frequency and duration) is a main determinant of the effect on glycemic control. However, a slightly older review that included 34 RCTs with a minimum treatment duration of 8 weeks reported that combined aerobic and resistance exercise improved glycemic control more than aerobic exercise alone (MD -0.67%, CI -0.93% to -0.40%; MD -0.60%, CI -0.98% to -0.27%, respectively).\textsuperscript{268}

Five reviews investigated only resistance exercise\textsuperscript{271,274,278} or aerobic exercise alone.\textsuperscript{270,273} All reported improved HbA1c for intervention groups compared with control. Four reviews examined other specific types of exercise. Of these, 3 reviews examined tai chi,\textsuperscript{272,279,280} and one investigated yoga\textsuperscript{74}; none found a statistically significant effect on glycemic index, weight, or cardiovascular risk.

**Dietary interventions.** Four eligible reviews on dietary interventions were restricted to RCTs and reported on both HbA1c and weight/BMI.\textsuperscript{105,281-283} Adverse events were not reported, and none reported gender distribution or sex effects. None originated from an organization known for high-quality systematic reviews. The number of trials ranged from 8 to 20.

The largest review (n=20 studies)\textsuperscript{105} examined 6-month outcomes after various dietary interventions (ie, high fiber, high protein, low carbohydrate, low glycemic index, Mediterranean,
vegan, or vegetarian) compared with one of multiple control diets (ie, ADA, high glycemic index, low fat, or low protein). This review concluded that the Mediterranean diet was the most effective at decreasing HbA1c (n=3; MD -0.41%, CI -0.58% to -0.24%) while low-carbohydrate (n=7; MD -0.12%), low glycemic index (n=3; MD -0.14%), and high protein (n=2; MD -0.28%) were all more effective than diets with higher carbohydrate content. Both the Mediterranean and low-carbohydrate diets produced weight loss (MD -1.84 and -0.69 kg, respectively), but the difference was statistically significant only for the Mediterranean diet.

**Quality improvement.** Three reviews evaluated different quality improvement strategies. None provided the gender distribution of included studies and none reported sex effects. The largest review (n=52 studies) examined the effect of care management on glycemic control. Statistically significant reductions in HbA1c were found (WMD -0.22%, CI -0.40% to -0.04%), but given high heterogeneity, the authors suggested these results were unlikely to be important. The second eligible review included 26 studies, which investigated 9 quality improvement strategies applied to rural settings. Quality improvement was more effective than control for reducing HbA1c (MD -0.41%, CI -0.75% to -0.07%), and subgroup analyses showed that using multiple strategies and a community setting for interventions were more effective compared with fewer strategies or a clinical setting, respectively. The third review (n=20 studies) was restricted to RCTs that evaluated computerized decision support systems (CDSS) in primary care. This qualitative synthesis reported that CDSS alone or with reminders improved process of care but did not affect health outcomes, whereas when CDSS was combined with feedback on self-management performance or case management, HbA1c improved.
APPENDIX G. OVERALL EFFECTS OF SELECTED INTERVENTIONS: CHRONIC PAIN

The information provided in this appendix reflects what was reported by the authors of the original, primary publication. We have not separately assessed the article for quality or accuracy.

Effects by Chronic Pain Condition

**Chronic low back pain.** Twelve reviews evaluated exercise for CLBP. One additional review compared acupuncture, spinal manipulation, and exercise. Exercise interventions included yoga and Pilates (n=6 studies), core stability, aerobic or general exercise, mixed programs, and nonspecific programs. The largest review (n=40) found that all types of exercise compared with minimal care improved pain (MD -4.83 [100-point scale], CI -9.36 to -0.30) and disability (MD -6.41, CI -9.76 to -3.05). The number of exercise sessions was significantly associated with increasing effect sizes on pain, suggesting that for each additional session, the effect size would increase by 0.13 (CI 0.02 to 0.24). Another review of similar size (n=39 studies) found that strength/resistance (n=11; ES: -0.50, CI -0.77 to -0.24) as well as coordination/stabilization-focused exercise programs are effective in reducing pain (n=12, ES -0.47, CI -0.77 to -0.18). Four reviews found moderate evidence of yoga or Pilates reducing pain over short-term, but not long-term, follow-up (largest review n=4; SMD -0.61, CI -0.97 to 0.26), while 2 reviews found no significant effect on pain or functionality (n=4) when comparing Pilates with no treatment or lumbar stabilization. Adverse events were rarely reported for exercise interventions.

Five reviews and one IPD meta-analysis evaluated either acupuncture or chiropractic manipulation compared with usual care, no treatment, or sham intervention. One additional review compared acupuncture, spinal manipulation, and exercise. The largest review of acupuncture (n=32 studies) reported significant reductions in self-reported pain (n=3-4; SMD -16.76 to -0.72; CI -33.33 to -0.19) and overall function (n=3-4; SMD -0.94 to -0.36; CI -1.61 to -0.04). However, significant, unexplained variation in treatment effects limit these findings. Another review found acupuncture to be more effective in reducing pain compared with sham acupuncture (n=5; SMD 0.20; CI 0.09 to 0.32) or no intervention (n=5; SMD 0.49; CI 0.33 to 0.64). The most recent large review of chiropractic manipulation (n=26 studies) reported significant but not clinically relevant effects on short-term pain relief (SMD -4.16; CI -6.97 to -1.36) and functional status (SMD -0.22; CI -0.36 to -0.07). Manual therapy was no more effective than sham intervention or exercise.

Two reviews evaluated injections (one on facet joint and the other on epidural), and using qualitative synthesis, found fair evidence for reduction in pain. Complications were rare. One additional large qualitative review examined multiple nonsurgical interventional therapies for low back pain, finding that epidural steroid injections were moderately effective for short-term but not long-term symptom relief. There was insufficient evidence for other interventional therapies for chronic low back pain, including facet joint injection and intradiscal steroid injection. Sex effects were not examined in any review.

One Cochrane review evaluated 30 trials of behavioral therapy compared with waitlist or usual care controls for CLBP. Operant behavior therapy was more effective than waitlist control (n=3; SMD -0.43, CI -0.75 to -0.11), and behavioral treatment was more effective than usual care.
for short-term pain relief (n=2; MD -5.18, CI -9.79 to -0.57). This review did not examine sex effects. Another review that evaluated many types of interventions (n=70) examined 21 behavioral studies and agreed with the conclusions of the Cochrane review. One review tested neurophysiological education only, but did not find clinically significant changes.

One industry-sponsored review of medications, using indirect meta-analysis, evaluated serotonin norepinephrine reuptake inhibitors (SNRIs) and SSRIs compared with NSAIDs, tramadol, and narcotics and found that, compared with placebo, scheduled opioids were most effective in reducing pain (n=8; SMD -0.39, CI -0.47 to -0.31), followed by nonscheduled opioids, Cox II inhibitors, and duloxetine.

One review of QI interventions evaluated pain rehabilitation programs compared with controls. This review identified 18 trials, finding a moderate effect of pain rehabilitation programs on improving disability (beta=0.429, SE=0.160, p=0.0009) and quality of life (beta=0.417, SE=0.033, p<.001).

**Fibromyalgia.** Of the 34 reviews of FM interventions, 11 were conducted by an organization known for high-quality reviews. Fourteen examined medication treatment for FM: 7 for antidepressants, 5 for anticonvulsant agents, and 2 for both classes of drugs. Of the reviews of different antidepressant drug classes, one examined a tricyclic agent, and 3 focused on SNRIs. A review that included 7 studies focusing specifically on FM found that amitriptyline reduced pain more than placebo with an RR of 2.9 (CI 1.7 to 4.9) based on 4 studies with second-tier data. Another review that included 10 studies focused on FM that SNRIs, specifically duloxetine and milnacipran, had small beneficial effects on both pain with a SMD -0.23 (95% CI -0.29 to -0.18), and quality of life with a SMD of -0.20 (95% CI of -0.25 to -0.14). For this review of SNRIs, the intended subgroup sex effects analyses for outcome of pain were not conducted due to lack of available individual patient-level data. This same review noted no difference between placebo and SNRI with respect to serious adverse effects, though more patients withdrew for adverse events in the SNRIs arms than in the placebo arm with an RR of 1.83 (CI 1.53 to 2.18). Additional reviews that focused solely on milnacipran, duloxetine, or a mixture of SNRIs, milnacipran, SSRIs, and tricyclic antidepressants supported the above findings related to pain relief.

The anticonvulsant agents gabapentin and pregabalin were found to improve pain in 5 reviews. In one trial of 150 patients included in the most recent of these reviews, gabapentin was noted to decrease pain more than 30% from baseline at 2400mg daily versus placebo with an RR of 1.6 (CI 1.07 to 2.42). Similarly, pregabalin reduced fibromyalgia pain 50% compared with placebo (RR 1.59, CI 1.33 to 1.90) based on high-quality evidence from 5 studies including 3256 patients, and an RR of 1.5 to 1.7 in a second review at doses 300mg to 600mg per day. A dose-response relationship with pregabalin was also noted for both pain control and adverse events, which were most frequently dizziness and somnolence. Serious adverse events were not higher with anticonvulsant treatment.

We identified 6 reviews with meta-analyses that included trials of various forms of exercise as treatment for pain associated with FM. Four of 5 types of exercise were found to decrease pain when compared with placebo, but it is less clear if any types of exercise are superior compared with other forms of exercise. Specifically, low- to moderate-quality evidence from 7 studies suggests that supervised aquatic exercise training causes moderate decreases in pain with
an MD of -6.59 (CI -10.71 to -2.48) on a 100-point scale; but this review did not find a significant improvement in pain compared with land-based exercise. Another review also found improvement with aquatic therapy, especially if the duration was at least 20 weeks. Aerobic exercise was also found to lead to significant reductions in pain with an SMD of -0.31 (CI -0.46 to -0.17). In a multiple treatment comparison analysis, low-quality evidence suggests that moderate to high-intensity resistance exercise training reduces pain in FM. Adverse effects related to aerobic exercise were infrequent.

Results for meditative movement therapies (MMT) such as yoga, tai chi, and qigong were conflicting. One review evaluating 7 studies found that these therapies did not reduce pain but may improve health-related quality of life. Another review found a medium-to-high effect size for MMT based on weak data. In both reviews, MMT was found to be safe with a low rate of dropout due to adverse events in one review and no serious adverse events in the second.

Six reviews examined different types of behavioral interventions. One review of 21 studies conducted in adults with FM found low-quality evidence that CBT compared with control produced an end-of-treatment benefit of 0.5 points on a scale from 0 to 10 with respect to pain (SMD -0.27, CI -0.47 to -0.07) and a benefit of 0.7 points for reducing disability (SMD -0.30, CI -0.51 to -0.08). Another review of 23 studies found that CBT produced a moderate positive effect on pain (n=8; Hedge’s g 0.60, CI +0.43 to +0.76). Psychological treatments other than CBT had a smaller effect size (n=18; Hedge’s g 0.27, CI +0.17 to +0.37). Mindfulness-based stress reduction also led to a small beneficial effect on FM-related pain and quality of life, based on generally low-quality evidence. The same was true of electromyogram-based biofeedback, but not electroencephalogram-based biofeedback. No specific information was noted about adverse events.

Five reviews examined acupoint stimulation/acupuncture for FM and came to similar conclusions. These reviews included from 9 to 16 RCTs. Compared with sham acupuncture, acupuncture did not significantly reduce FM-related pain. Based on one RCT with 16 participants, acupuncture did appear to be more effective than no treatment with an RR of -30.19% (CI -55.23% to -5.15%) and more effective as an adjunct to standard therapy based on one RCT with 58 patients with an RR of -37.50% (CI -48.75% to -26.25%). There was some low-quality evidence that acupuncture is more effective at reducing pain than medications. Electro-acupuncture may be better than manual acupuncture at pain reduction. That review concluded there is insufficient information about adverse events to draw definitive conclusions; however, another review reported adverse effects from 8 of 25 studies as bruising, nausea, fainting, discomfort with needles, and temporary edema of the hand. One trial reported that a patient had mild scalding on the skin after being included in a cupping group. One review examined 4 studies of chiropractic care for FM and concluded that there was no significant difference between intervention and control groups.

Knee OA. We identified 8 reviews that examined intervention effects on pain due to knee OA. One large review evaluating 54 trials found high-quality evidence that land-based therapeutic exercise significantly reduced pain compared with control (nonexercise or nontreatment) with an SMD -0.49 (CI -0.39 to -0.59). Eight studies included measurement of serious adverse events, and none were reported. Similar, a review of pre-operative interventions included 4 trials which found that exercise programs significantly decreased pain from knee OA.
with a SMD -0.43 (CI -0.13 to 0.73). One review of 7 trials of proprioceptive-based exercise found no improvement in pain and a small improvement in function.\textsuperscript{303}

One review of medications used a multiple treatment comparison meta-analysis to examine the effects of topical NSAIDs for pain in patients with chronic musculoskeletal pain.\textsuperscript{166} The review included 34 studies, 16 of which focused on knee OA. Topical NSAIDs were more effective than placebo (RR 1.29 for 50% pain reduction or equivalent measure; CI 1.21 to 1.38) but not more effective than oral NSAIDs for pain in the context of osteoarthritis (RR 1.02, CI 0.94 to 1.11). Topical NSAIDs led to fewer gastrointestinal side effects than oral NSAIDs but led to an increase in local adverse events (mostly skin-related). This review did not evaluate sex effects.

Three reviews examined injections for knee OA. One review\textsuperscript{165} examined the trajectory of intra-articular hyaluronic acid therapeutic effects on knee OA and found a beneficial effect on pain that peaked at 8 weeks post-injection (n=54; ES 0.46, CI 0.28 to 0.65), diminishing to a small effect at 24 weeks (ES 0.21, CI 0.10 to 0.31). One review\textsuperscript{172} considered 4 RCTs and 2 studies that compared platelet-rich plasma (PRP) injections as compared with intra-articular normal saline or hyaluronic acid. Based on 4 studies, this review found that PRP led to greater improvements in the WOMAC, a composite score\textsuperscript{304} of pain, stiffness and function with an MD of -18.0 (-28.8 to -8.3). Of note, they report that there was an increase in nonspecific adverse events among those who received PRP versus control (8.4% vs 3.8%). A third review of 6 RCTs found no benefit of joint lavage on pain due to knee OA.\textsuperscript{169} None of these reviews evaluated sex effects.

One review using IPD meta-analysis of 31 studies examined acupuncture for chronic pain including 9 studies of knee OA.\textsuperscript{168} Acupuncture decreased pain of knee OA compared with both no acupuncture (SMD 0.57, CI 0.29 to 0.85) and sham acupuncture (SMD 0.37, CI 0.03 to 0.72). This review did not evaluate sex effects.
APPENDIX H. SYSTEMATIC REVIEW REFERENCES BY CONDITION

This appendix lists full citations for included systematic reviews in alphabetical order by condition. The asterisk beside a reference indicates the review reported sex effects.

**Depressive Disorders (86)**


Cameron IM, Reid IC, MacGillivray SA. Efficacy and tolerability of antidepressants for sub-threshold depression and for mild major depressive disorder. J Affect Disord 2014;166:48-58. PMID: 25012410.


Type 2 Diabetes (114)


Evidence-based Synthesis Program


Mapping the Evidence: Sex Effects in High-Impact Conditions for Women Veterans


Mapping the Evidence: Sex Effects in High-Impact Conditions for Women Veterans


Chronic Pain (68)


