

Accepted author's manuscript. Published in final edited form as: Journal of the American Geriatrics Society
JAGS 2020 (in press). Publisher DOI: [10.1111/jgs.16697](https://doi.org/10.1111/jgs.16697)

1

2 **Outcome measures for interventions to reduce inappropriate chronic drugs: a narrative review**

3

4

5 Carole E. Aubert, MD,^{1,2,3,4} Eve A. Kerr, MD, MPH,^{3,4,5,6} Jennifer K. Maratt,^{7,8} MD, MS, Mandi L. Klamerus,
6 MPH,³ Timothy P. Hofer, MD, MSc.^{3,4,5,6}

7

8 ¹ Department of General Internal Medicine, Bern University Hospital, University of Bern, Bern, Switzerland;

9 ² Institute of Primary Health Care (BIHAM), University of Bern, Switzerland;

10 ³ Veterans Affairs Center for Clinical Management Research, Ann Arbor, MI, USA;

11 ⁴ Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, USA;

12 ⁵ Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, MI, USA;

13 ⁶ Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA;

14 ⁷ Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA;

15 ⁸ Richard L. Roudebush Veterans Affairs Medical Center, Indianapolis, IN, USA.

16

17 **E-mail addresses:** caubert@umich.edu; ekerr@umich.edu; mandi.klamerus@va.gov; jmaratt@iu.edu;
18 thofer@umich.edu

19

20 **Running title:** Deprescribing intervention measures review.

21

22 **Manuscript category:** Narrative review.

23

24 **Word count:** 3,595

Abstract word count: 281

25 **Number of Tables:** 3

Number of Figures: 1

26 **Number of Supplements:** 5

Number of References: 42

27

28 **Corresponding author:** Carole E. Aubert, MD, North Campus Research Complex Building 16, 2800 Plymouth
29 Rd, Ann Arbor MI 48109-2800; +1 734 845 3504; caubert@umich.edu, @aubert_carole

30

31 **Funding:** Dr. Aubert was supported by an Early Postdoc.Mobility grant from the Swiss National Foundation.

31

32 **STRUCTURED ABSTRACT**

33 **Background:** Inappropriate prescribing is a highly important problem, given the growing aging
34 multimorbid population with associated polypharmacy. An increasing number of studies have recently
35 developed and tested interventions to withdraw inappropriate drugs, a process called deprescribing.
36 However, we still lack complete information on the types and prevalence of measures used to assess the
37 success of such interventions.

38 **Objective:** To categorize and synthesize the full spectrum of measures used in intervention studies
39 focused on reducing inappropriate prescribing of chronic drugs in adults, in order to standardize
40 measurements in future studies and help researchers design studies inclusive of the important measure
41 types.

42 **Design:** We searched Ovid/MEDLINE to identify intervention studies focused on deprescribing chronic
43 drugs in adults, published between 2010 and 2019.

44 **Measurements:** We extracted data on study characteristics, intervention components, and outcome
45 measures. We categorized and synthesized the measures using a comprehensive and systematic
46 framework, separating measures of intended and unintended consequences.

47 **Results:** Most (90/93) studies used measures of appropriate prescribing, such as drug cessation or dose
48 reduction. The following measures were used infrequently across studies: patient-reported experience,
49 preferences, and outcome (12 (13%), 2 (2%), and 25 (27%) studies, respectively); provider-reported
50 experience (11 (12%) studies); patient-provider interaction (4 (4%) studies); and measures of unintended
51 consequences (24 (26%) studies). Studies varied in the type and number of measures assessed, ranging
52 from 1 to 20 different measures by study.

53 **Conclusion:** To ensure initiation, success, and long-term sustainability of deprescribing, it is important
54 to assess the success of intervention studies using clinically relevant patient- and provider-centered
55 measures. This categorized synthesis of outcome measures used in deprescribing studies may facilitate
56 implementation of important measure types (e.g., patient reported measures, measures of unintended
57 consequences) in future studies.

58 **Key words:** deprescribing; inappropriate medication; withdrawal; interventions; measures.

59 INTRODUCTION

60

61 Up to 30% of medical services are considered low-value, i.e., may result in more harm than benefit.¹⁻³
62 Inappropriate prescribing is increasingly seen among the growing older multimorbid population,^{4,5} with
63 up to one-third receiving inappropriate prescriptions.⁶ In response, the Choosing Wisely initiative
64 regularly publishes recommendations to minimize low-value prescribing.¹ While an increasing number
65 of interventions focused on deprescribing inappropriate medications,⁷ deprescribing chronic medications
66 remains a complex process associated with barriers at both patient and provider levels,^{8,9} particularly for
67 medications, whose use was prompted by unpleasant symptoms. Fear of worsening symptoms may lead
68 to resistance towards stopping these medications.¹⁰ Further, clinicians lack time and resources for
69 deprescribing, report low self-efficacy for stopping therapy, and feel uncertain about clinical
70 consequences of deprescribing (e.g., stroke following antihypertensive drug reduction).¹¹ To ensure
71 feasibility and sustainability of deprescribing, intervention studies should assess not only whether a
72 medication was stopped or the dose reduced, but also patient-relevant clinical outcomes and patient and
73 provider experience and preferences. The measures should capture both intended effects and unintended
74 harms, a key priority identified by Choosing Wisely and patient advocates.^{12,13} However, deprescribing
75 intervention studies have highly variable outcome measures and rarely include clinical outcomes, as
76 outlined in two reviews in older adults.^{8,14} These reviews did not detail the types and frequency of use of
77 the different measures, and only assessed controlled trials.^{8,14} This global paucity of clinical outcomes
78 and heterogeneity of measures may be explained by a lack of guidance. It is also more challenging to
79 collect information on experience, preferences and clinical outcome measures, as this requires longer
80 follow-up periods, prospective designs, and broader expertise.

81 We recently reviewed the literature to characterize measures employed in 117 interventions to reduce
82 low-value care.¹⁵ We found that measures focused largely on utilization and rarely addressed patient-
83 centered outcomes or unintended consequences. The search strategy was not tailored to identify low-
84 value prescribing of chronic medications and included only 44 studies focused on prescribing for

85 predominantly acute medications (two-thirds addressed acute antibiotic use). Given the unique
86 challenges of stopping chronic medications, the measures to assess the impact of interventions may be
87 notably different from those used in studies focused on stopping acute medications.
88 Based on this review, we suspected that outcome measures reported across deprescribing intervention
89 studies for chronic medications would also lack coverage of important measure types.¹⁵ Given the lack
90 of prior reviews, and the need to standardize outcome measures for further studies,¹⁶ we sought to
91 provide the first review to: 1) identify measures used in recent studies evaluating the effect of
92 interventions to reduce inappropriate prescribing of chronic medications in adults, including prescribing
93 practices, clinical outcomes, cost/value, and patients' and providers' experience and interaction, and 2)
94 categorize and synthesize these measures, using a comprehensive systematic framework, to provide
95 deprescribing study designers with a list of candidate measures within each category.

96

97 **METHODS**

98

99 **Search strategy**

100 We performed a literature search in Ovid/MEDLINE search from January 1, 2010, to October 13, 2019
101 to identify original studies of any design reporting outcome measures of interventions to reduce
102 inappropriate prescribing of chronic drugs in adults (**Supplementary Text S1**). A separate search
103 strategy was used for benzodiazepine-related drugs, without the term “appropriate prescribing” given
104 that most use is considered inappropriate. The search was restricted to Ovid/MEDLINE, as we felt that
105 this source alone would be sufficient to identify articles that would allow us to capture the full spectrum
106 of available measures. Inclusion criteria were: adult population; original study (i.e., not a review or
107 meta-analysis); intervention to reduce the use of a least one chronic inappropriate drug. We included
108 both quantitative and qualitative studies. We excluded studies that focused on: 1) only new drug
109 prescriptions (e.g., new prescription of proton pump inhibitor during hospitalization) or only on short-
110 term or acute drugs (e.g., antibiotic for urinary tract infection); we didn't use a clear cut-off to define a

111 drug as non-chronic, as it varied depending on the drug class; 2) reducing polypharmacy in general
112 without assessing prescribing appropriateness; 3) deprescribing as part of a global intervention not
113 focused on reducing inappropriate prescribing; 4) inappropriate prescribing assessed globally as
114 potentially inappropriate prescription, potential prescribing omission, inappropriate dosage or drug
115 interactions. We focused on interventions to deprescribe chronic drugs, because the specific challenges
116 and barriers are likely to be different than those for prescribing acute drugs or new drugs.

117

118 **Measure definition and categorization**

119 A measure was defined as any assessment of prescribing practice, clinical outcome, cost/value, or
120 experience following the deprescribing intervention. We classified the measures used in the studies into
121 several categories, adapted from a framework previously developed by our research team
122 **(Supplementary Table S1):**¹⁵ 1) measure specification (count, scale, proportion); 2) measure type
123 (appropriateness, utilization/ordering, intermediate outcome, outcome, patient-reported outcome
124 (PROM), patient-reported experience (PREM), patient preferences, provider-reported experience,
125 patient-provider interaction, cost-related); 3) measure reporting type (patient, provider,
126 medical/pharmacy record, validated scale/questionnaire, non-validated scale/questionnaire, blinded
127 assessment); 4) measure of unintended consequence (including substitution of an alternative low-value
128 drug, underuse of the drug being intervened upon, underuse of related services, PREM, provider-
129 reported experience, patient-provider interaction, patient selection, care location shift, harmful outcome,
130 reimbursement), which were classified as “definite” if the study specifically reported it as such in the
131 methods section, or “possible” if it was inferred by the reviewer. Appropriateness and
132 utilization/ordering measures were further classified into subcategories: cessation, dose reduction, new
133 prescription, switch for another drug. Utilization/ordering measures included prescribing measures not
134 assessing the appropriateness of the drug.

135

136 **Data extraction**

137 The first author (CEA) performed the literature search and used a standardized form to extract relevant
138 data. Data on study characteristics included first author name, publication year, design, setting,
139 participants (with specific inclusion criteria such as older age, multimorbidity, polypharmacy), number
140 and class(es) of drug(s), and intervention aim, target (patient or provider), description and type (e.g.,
141 education, feedback, drug review). Data on measures included information required for categorization.

143 **Data analyses**

144 Separate articles referring to the same study were grouped for analysis. Similar measures across these
145 articles were also merged. We present study characteristics as frequencies/percentage of studies (number
146 of studies with characteristic relative to total number of studies), and measures as
147 frequencies/percentage of measures (number of measures of a specific type relative to total number of
148 measures) and percentage of studies, respectively. We summarized all measures used in the studies,
149 grouping similar measures (e.g., drug cessation, intervention acceptance) used across different studies,
150 to provide a synthesized reference list of potential measures to consider in future deprescribing studies.

152 **RESULTS**

154 **Studies included**

155 From the 4,190 articles identified in Ovid/MEDLINE, 4,041 were excluded upon review of the title
156 and/or abstract (**Figure 1**). Of the remaining 149 articles, 44 were excluded upon review of the full-text,
157 resulting in 105 articles included in the review. Eight studies published their results through two to four
158 separate articles, so that the total of 105 articles represents 93 unique studies. A complete list of the 105
159 articles is provided in **Supplementary Text S2**.

161 **Study population, setting, design and drug classes**

162 Most of the 93 studies (n=60, 65%) focused on older patients. Fifty-one (55%) studies were conducted
163 in the outpatient setting, 27 (29%) in long-term care, 19 (20%) in the inpatient setting, and 8 (9%) in the
164 pharmacy (**Table 1**). A control group was used in 42 (45%) studies, of which half employed
165 randomization. The most frequent drug classes studied were sedative-hypnotics (in 64 (69%) studies)
166 and antipsychotics (in 43 (46%) studies). Forty-two (45%) studies involved a single drug class. Study
167 characteristics are detailed in **Supplementary Table S2**.

168

169 **Intervention characteristics**

170 The interventions were most often multifaceted and targeted a patient (in 44 (47%) studies) and/or a
171 provider (in 85 (91%) studies). The most frequent intervention types were a review of drug
172 appropriateness and indication in 40 (43%) studies, followed by education at the patient or provider
173 level in 29 (31%) and 31 (33%) studies, respectively. The intervention types used in each study are
174 detailed in **Supplementary Table S2**.

175

176 **Outcome measures characteristics within studies**

177 Across the 93 studies, we identified 511 outcome measures. We present frequencies of each measure
178 type in **Table 2**. Complete drug cessation was the most frequently assessed measure, in 79 (85%)
179 studies. Thirty-two (34%) studies used at least one patient-reported measure, including PROMs,
180 PREMs, and patient preferences. One fourth of the studies (n=24) reported using at least one measure of
181 unintended consequences (e.g., withdrawal symptoms or use of restraints for agitation). Non-patient
182 reported outcome measures (e.g., hospitalizations), including intermediate outcomes (e.g., uptake of
183 deprescribing intervention by the prescribing physician), were used in 46 (49%) studies. Provider-
184 reported experience, patient-provider interaction, and cost-related measures were rarely used. **Table 3**
185 provides a synthesized and categorized list of all measures used across the studies, with some examples.
186 The frequencies and types of measures used in each study are listed in **Supplementary Table S3**.

187

188 **Outcome measures source within studies**

189 We present frequencies of each measure source (i.e., patient-reported, provider-reported,
190 medical/pharmacy record, validated/non-validated scale or questionnaire, blinded assessment) in **Table**
191 **2**. Medical or pharmacy records were the most frequent sources used for measures (86 (93%) studies).
192 Blinded measures assessment was performed in only 11 (12%) studies (50% of the randomized trials).

194 *Appropriateness and utilization/ordering measures*

195 Thirty-four (37%) studies used both appropriateness and utilization/ordering measures (i.e., without
196 assessing appropriateness of prescribing), while 56 studies (60%) measured only appropriateness, and a
197 single study (1%) only utilization/ordering. Appropriateness and utilization measures included cessation,
198 dose reduction, new prescription, and switch for another drug, either alone or in combination. For
199 example, Ailabouni et al. evaluated the number of drugs prescribed (utilization/ordering measure) and
200 the Drug Burden Index (appropriateness measure), while Brodaty et al. assessed cessation of
201 inappropriate antipsychotics (appropriateness measure) and prescription rate of other psychotropic drugs
202 (utilization/ordering measure).^{17,18} Studies assessing several drug classes most often reported these
203 measures for all classes combined and for each class separately. For example, Ammerman et al. assessed
204 discontinuation rate of any potentially inappropriate medication evaluated, as well as discontinuation
205 rate of anticholinergics, nonsteroidal anti-inflammatory drugs, proton pump inhibitors, peripheral alpha
206 blockers, benzodiazepines, antihistamines, and antipsychotics separately.¹⁹

208 *Patient-reported measures*

209 Twenty-five studies (27%) used PROMs, while only 12 (13%) and 2 (2%) studies assessed PREMs and
210 patient preferences, respectively. PROMs mostly included quality of life or perceived health status, as
211 well as drug-specific outcomes, such as sleep quality, drug dependence, cognition, sedative side effects
212 or withdrawal/anxiety/depression symptoms for sedative-hypnotics, or gastrointestinal symptoms for
213 proton pump inhibitors. PREMs most often evaluated a patient's experience with the intervention (e.g.,

214 satisfaction with educational material) or of the tapering process (e.g., reasons for tapering difficulties).
215 Patient preferences measures included reasons for refusing deprescribing or preferences for the
216 intervention.

217

218 *Provider-reported experience and patient-provider interaction measures*

219 Eleven (12%) studies evaluated provider-reported experience measures, including experience,
220 satisfaction or acceptance of the intervention, as well as self-efficacy for deprescribing. Only 4 (4%)
221 studies used patient-provider interaction measures, reporting the number of counseling occasions,
222 personal interactions, discussion documentation, and drug review with the patient.

223

224 *Non-patient reported intermediate outcome and outcome measures*

225 Thirty-three (35%) and 19 (20%) studies included a non-patient-reported outcome or intermediate
226 outcome measure, respectively. Intermediate outcome measures often related to acceptance rate of
227 deprescribing recommendations. Outcome measures included healthcare services utilization
228 (hospitalization, length of stay, ambulatory visits) and mortality. Additionally, outcome measures often
229 included outcomes related to specific drugs (e.g., falls or confusion for sedative-hypnotics,
230 neuropsychiatric symptoms or use of a seclusion room for antipsychotics, incidence of cardiovascular
231 events for antihypertensive and lipid-lowering drugs).

232

233 *Cost-related measures*

234 Ten (11%) studies assessed effects on costs. The majority of these measured drug costs, while three
235 (3%) evaluated the cost of the intervention (e.g., provision of educational material) and two measured
236 the cost of healthcare services utilization. Only two (2%) studies used a value measure, specifically
237 assessing cost-utility of the intervention.

238

239 *Qualitative measures*

240 While all studies used quantitative measures, only 18 (19%) also performed a qualitative assessment.
241 Qualitative measures included patient and provider experience, acceptance or satisfaction with the
242 intervention assessed qualitatively (e.g., by interview), key messages remembered by providers, reasons
243 for not deprescribing or for restarting a deprescribed drug, feasibility of the intervention, patient
244 perception of deprescribed drugs, physician impression of deprescribing rounds, communication
245 preferences, or decisions during discussions between patients and providers.

246

247 *Measures of unintended consequences*

248 Twenty-four (26%) studies reported at least one measure of unintended consequences, which
249 represented 10% (n=52/511) of all measures. Among them, 21 were clearly mentioned as such in the
250 methods, and thus classified as “definite,” while 31 were considered as unintended consequences by the
251 reviewer and classified as “possible.” Unintended consequences included changes in symptoms or
252 withdrawal related to drug tapering, use of restraints or substitute drugs, changes in laboratory
253 parameters, as well as adverse events during deprescribing, such as hospitalization, falls, death or
254 cardiovascular events. Of the 52 measures, outcome measures documenting unintended consequences
255 were the most frequent (n=21, 40%), followed by PROMs (n=15, 29%), utilization/ordering measures
256 (n=10, 19%), appropriateness measures (n=5, 10%) and provider-reported experience measures (n=1,
257 2%).

258

259

260 **DISCUSSION**

261

262 In this review of 93 deprescribing studies, we found that almost all authors used an appropriateness
263 measure assessing change in prescribing, most frequently drug cessation, to examine the impact of their
264 interventions. Less often they simply used a measure of utilization or ordering, without taking into
265 account appropriateness of medication indication and/or dosage. Less than half of the studies examined

266 non patient-reported outcomes, such as mortality or utilization of healthcare services. Patient-provider
267 interaction, provider-reported experience, and cost-related measures were used infrequently and only
268 26% of the studies evaluated unintended consequences of deprescribing.

269 Outcome measures were uncommon and inconsistently used across all studies. Not surprisingly, any
270 specific measure employed was usually related to the type of intervention. For example, studies on
271 sedative-hypnotic drugs evaluated the incidence of falls or the use of other psychotropic drugs, while
272 studies on proton pump inhibitors assessed rebound dyspeptic symptoms or the use of a rescue drug
273 such as a H2 blocker. Interventions with a strong focus on the patients were more likely to assess
274 patient-reported measures, although these were present in less than one third of the studies, and
275 measures of patient experience and preferences were particularly rare.

276 The literature suggests that deprescribing is more likely to be successful when individual patient
277 context, preferences, and goals are considered,²⁰⁻²² particularly when patients may have withdrawal
278 symptoms , such as for psychotropic drugs or proton pump inhibitors,^{23,24} and thus education and active
279 participation for self-management is required.

280 Although a strong focus on patient involvement is important, deprescribing remains most often initiated,
281 directed, and sometimes required by providers, who may face multiple barriers,¹¹ so studies should also
282 assess the experience of the providers with the interventions. However, only a minority of authors
283 employed provider-reported experience measures, while four studies assessed patient-provider
284 interactions, including shared-decision making. For example, Carr et al. assessed the number of
285 conversations around benzodiazepine cessation, and found that patients with more conversations had
286 higher rates of deprescribing.²⁵ Deprescribing chronic drugs may lead patients to fear or even experience
287 withdrawal symptoms. Thus, it is important that providers understand how the patients experience
288 potential harms and benefits of reducing the drugs, and discuss and implement deprescribing in a
289 shared-decision-making process, a key facilitator to deprescribing.²⁶ Future studies should more
290 consistently assess provider experience and patient-provider interactions. Tools such as CollaboRATE

291 or the revised Patients' Attitudes Towards Deprescribing questionnaire could be used for this
292 purpose.^{27,28}

293 Specific barriers and facilitators for deprescribing were largely assessed by qualitative studies, mostly
294 by interviewing or surveying patients or providers, while qualitative methods were rarely used in
295 intervention deprescribing studies (only 18 of the 93 (19%) studies included in this review).^{21,29-33}

296 Qualitative research requires particular expertise and resources that differ from purely quantitative
297 methods,³⁴ but allows a broader assessment of barriers and facilitators, as well as patient- and provider-
298 reported experiences than quantitative measurement alone, so that it should be integrated in
299 deprescribing intervention studies.³⁵

300 Withdrawing medications is recommended when harms outweigh benefits.⁷ However, deprescribing
301 may result in withdrawal symptoms (e.g., sweating or irritability for benzodiazepines), return of the
302 medical condition (e.g., heartburn for proton pump inhibitors), increased use of healthcare services, or
303 incidence of a new condition precluded after a preventive medication is reduced (e.g., stroke for
304 antihypertensive medications).³⁶ It is therefore important to carefully monitor the patients during and
305 after the deprescribing process, and to measure potentially unintended consequences, such as more
306 frequent than expected new or recurrent symptoms, or higher healthcare services utilization.¹³ Our
307 review suggests an important gap in this context, since only 27% and 35% of the authors assessed
308 patient-reported and other outcome measures, respectively, and one fourth assessed unintended
309 consequences of the interventions. Finally, since some of these outcomes are infrequent or may occur
310 only after a relatively long follow-up period, it is important to design the studies for these outcomes if
311 important clinically. In our review, only one fourth of the interventions were randomized, with blinded
312 measure assessment in only half of the randomized trials.

313 We found very little overlap in the number and types of outcome measures used across the studies.
314 Research on deprescribing will have little cumulative impact on patient care without a standardized
315 outcome set that covers the important types relevant to deprescribing. The lack of consistency in
316 outcome measures reported may be related to a lack of exemplars in the literature on which to base the

317 design of deprescribing intervention studies and the relatively recent interest in the topic. There were
318 indeed some initial attempts to develop outcome sets in the context of deprescribing, but these focused
319 on older patients with polypharmacy and on medication appropriateness more broadly.^{37,38} Thus, the
320 results may not be generalizable to other populations or to specific medications. For example, in those
321 studies, PROMs included cognitive functioning, patient perception of medication burden, and pain
322 relief. Those outcome measures may be particularly pertinent for older multimorbid patients with
323 polypharmacy, but less relevant for younger patients trying to stop proton-pump inhibitors, for example.
324 Outcome sets for older adults also have a strong focus on medication-related outcomes, such as therapy
325 duplication, complexity or adherence, all of which are related to polypharmacy. We did not limit our
326 work to older or multimorbid patients with polypharmacy and used a framework to develop a broader
327 but nonetheless synthesized set of measures for each category. This framework may serve any
328 deprescribing intervention study and help to ensure that relevant measures across the whole spectrum,
329 including patient- and provider-centered and unintended consequences measures, are included.
330 We found little consistency not only in the number and types of measures considered, but also in the
331 designs and intervention types of the studies. All these issues are important to ensure the success of
332 deprescribing interventions. The following criteria may serve as exemplars for future researchers: 1)
333 high evidence-based design (randomized controlled trial); 2) intervention component targeting not only
334 the providers, but also patients; 3) broad set of measures to assess the success and acceptability of
335 deprescribing, with both qualitative and quantitative assessment; and 4) follow-up period long enough to
336 evaluate sustainability of deprescribing, which may provide information on scalability. The OPTI-
337 SCRIPT Study (articles numbers 2-5 in **Supplementary Table S2** and **Supplementary Table S3**),³⁹⁻⁴² a
338 cluster randomized controlled trial conducted in an outpatient general care setting to deprescribe
339 multiple potentially inappropriate drugs, is such an exemplar. The feasible intervention targeted
340 providers (web-based algorithm, education, drug review) and patients (educational leaflets), and the
341 authors assessed not only prescribing practices, but also clinical outcome, patient-reported experience
342 and outcomes, provider-reported experience, and patient-provider interaction, using a mixed-method

343 process. In addition, patients were followed-up for 12 months and cost-utility and cost-effectiveness
344 were evaluated.

345 There are several limitations to this review. First, we did not grade the quality of the studies, because we
346 focused on outcome measures and not on the effectiveness of the interventions themselves. Nonetheless,
347 it is noteworthy that a minority of the studies were randomized and only 45% included a control group.
348 Second, we searched only Ovid/MEDLINE. However, this search identified a large number of articles,
349 and extending the search to other databases (e.g., EMBASE) did not significantly increase the number of
350 relevant articles. Third, we did not review unpublished or ongoing studies, and it is possible, although
351 unlikely, that ongoing studies are using a larger spectrum of measures. Our study also has several
352 strengths. First, we used a broad search strategy, including specific search terms to capture interventions
353 targeting the most frequent inappropriate drugs. This strategy was developed with a medical librarian
354 and tested for identification of the most relevant articles. Second, we used a comprehensive and
355 systematic categorization framework to capture a broad range of measures, including both intended and
356 unintended consequences of the interventions. Finally, we synthesized and categorized the measures to
357 help designers of future deprescribing intervention studies have access to the full spectrum of available
358 measures.

359 In conclusion, this review confirmed our hypotheses that the success of deprescribing is most
360 consistently evaluated by drug cessation or dose reduction, while patient- and provider-reported
361 experience, preferences and outcomes, as well as measures of unintended consequences, are
362 infrequently considered. To ensure success and sustainability of deprescribing, it is important that
363 intervention studies include measures that are more clinically meaningful and centered on patients and
364 providers. To allow assessment of rare outcomes and in-depth evaluation of patient and provider
365 preferences and experience, we suggest using a mixed-methods approach, combining a randomized
366 controlled design with qualitative and implementation assessments. Finally, to facilitate incorporation of
367 a broad spectrum of measures into those future studies, the synthesis and categorization of the available
368 measures and identified gaps offers a first reference list of measures that can be useful for any

369 deprescribing study. Further validation of these measures by patients and providers concerned by
370 inappropriate prescribing will ensure that measures relevant to the stakeholders are included in the
371 process of deprescribing.

372

373 **ACKNOWLEDGEMENTS**

374 We would like to thank Judith Ellen Smith from the Taubman Health Sciences Library of the University
375 of Michigan for her help in constructing the literature search.

376 **Conflicts of Interest**

377 The authors declare that they do not have a conflict of interest.

378 **Author Contributions**

379 CEA, EAK and TH designed the study. CEA conducted the literature review, extracted the data,
380 performed the analyses, interpreted the results and wrote the manuscript. MLK developed the
381 abstraction database for data abstraction. EAK and TH contributed to interpretation of the data. EAK,
382 TH, MLK and JM revised the manuscript critically for important intellectual content. All authors agreed
383 for submission of the final version of the manuscript.

384 **Sponsor's Role**

385 CEA was supported by a grant from the Swiss National Foundation which had no role in the study.

386

387

388 **REFERENCES**

- 389 1. American Board of Internal Medicine. Choosing Wisely: An Initiative of the ABIM Foundation.
390 Available at <http://www.choosingwisely.org>.
- 391 2. Brownlee S, Chalkidou K, Doust J, et al. Evidence for overuse of medical services around the
392 world. *Lancet* 2017;390:156-68.
- 393 3. de Vries EF, Struijs JN, Heijink R, Hendriks RJ, Baan CA. Are low-value care measures up to the
394 task? A systematic review of the literature. *BMC Health Serv Res* 2016;16:405.
- 395 4. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity
396 and implications for health care, research, and medical education: a cross-sectional study. *Lancet*
397 2012;380:37-43.
- 398 5. Yarnall AJ, Sayer AA, Clegg A, Rockwood K, Parker S, Hindle JV. New horizons in
399 multimorbidity in older adults. *Age Ageing* 2017;46:882-8.
- 400 6. Bradley MC, Motterlini N, Padmanabhan S, et al. Potentially inappropriate prescribing among
401 older people in the United Kingdom. *BMC Geriatr* 2014;14:72.
- 402 7. Reeve E, Gnjjidic D, Long J, Hilmer S. A systematic review of the emerging definition of
403 'deprescribing' with network analysis: implications for future research and clinical practice. *Br J Clin*
404 *Pharmacol* 2015;80:1254-68.
- 405 8. Gnjjidic D, Le Couteur DG, Kouladjian L, Hilmer SN. Deprescribing trials: methods to reduce
406 polypharmacy and the impact on prescribing and clinical outcomes. *Clin Geriatr Med* 2012;28:237-53.
- 407 9. Reeve E, Thompson W, Farrell B. Deprescribing: A narrative review of the evidence and practical
408 recommendations for recognizing opportunities and taking action. *European journal of internal medicine*
409 2017;38:3-11.
- 410 10. Reeve E, To J, Hendrix I, Shakib S, Roberts MS, Wiese MD. Patient barriers to and enablers of
411 deprescribing: a systematic review. *Drugs & aging* 2013;30:793-807.

- 412 11. Anderson K, Stowasser D, Freeman C, Scott I. Prescriber barriers and enablers to minimising
413 potentially inappropriate medications in adults: a systematic review and thematic synthesis. *BMJ open*
414 2014;4:e006544.
- 415 12. Thompson W, Reeve E, Moriarty F, et al. Deprescribing: Future directions for research. *Research*
416 *in social & administrative pharmacy : RSAP* 2019;15:801-5.
- 417 13. Bhatia RS, Levinson W, Shortt S, et al. Measuring the effect of Choosing Wisely: an integrated
418 framework to assess campaign impact on low-value care. *BMJ Qual Saf* 2015;24:523-31.
- 419 14. Thillainadesan J, Gnjdic D, Green S, Hilmer SN. Impact of Deprescribing Interventions in Older
420 Hospitalised Patients on Prescribing and Clinical Outcomes: A Systematic Review of Randomised Trials.
421 *Drugs & aging* 2018;35:303-19.
- 422 15. Maratt JK, Kerr EA, Klamerus ML, et al. Measures Used to Assess the Impact of Interventions to
423 Reduce Low-Value Care: a Systematic Review. *J Gen Intern Med* 2019;34:1857-64.
- 424 16. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials*
425 2017;18:280.
- 426 17. Ailabouni N, Mangin D, Nishtala PS. Deprescribing anticholinergic and sedative medicines:
427 protocol for a Feasibility Trial (DEFEAT-polypharmacy) in residential aged care facilities. *BMJ open*
428 2017;7:e013800.
- 429 18. Brodaty H, Aerts L, Harrison F, et al. Antipsychotic Deprescription for Older Adults in Long-term
430 Care: The HALT Study. *Journal of the American Medical Directors Association* 2018;19:592-600.e7.
- 431 19. Ammerman CA, Simpkins BA, Warman N, Downs TN. Potentially Inappropriate Medications in
432 Older Adults: Deprescribing with a Clinical Pharmacist. *Journal of the American Geriatrics Society*
433 2019;67:115-8.
- 434 20. Todd A, Jansen J, Colvin J, McLachlan AJ. The deprescribing rainbow: a conceptual framework
435 highlighting the importance of patient context when stopping medication in older people. *BMC geriatrics*
436 2018;18:295.

- 437 21. Weir K, Nickel B, Naganathan V, et al. Decision-Making Preferences and Deprescribing:
438 Perspectives of Older Adults and Companions About Their Medicines. *The journals of gerontology Series*
439 *B, Psychological sciences and social sciences* 2018;73:e98-e107.
- 440 22. Ostini R, Jackson C, Hegney D, Tett SE. How is medication prescribing ceased? A systematic
441 review. *Med Care* 2011;49:24-36.
- 442 23. Pottie K, Thompson W, Davies S, et al. Deprescribing benzodiazepine receptor agonists:
443 Evidence-based clinical practice guideline. *Canadian family physician Medecin de famille canadien*
444 2018;64:339-51.
- 445 24. Tannenbaum C, Martin P, Tamblyn R, Benedetti A, Ahmed S. Reduction of inappropriate
446 benzodiazepine prescriptions among older adults through direct patient education: the EMPOWER cluster
447 randomized trial. *JAMA internal medicine* 2014;174:890-8.
- 448 25. Carr F, Tian P, Chow J, et al. Deprescribing benzodiazepines among hospitalised older adults:
449 quality improvement initiative. *BMJ open quality* 2019;8:e000539.
- 450 26. Jansen J, Naganathan V, Carter SM, et al. Too much medicine in older people? Deprescribing
451 through shared decision making. *Bmj* 2016;353:i2893.
- 452 27. Forcino RC, Barr PJ, O'Malley AJ, et al. Using CollaboRATE, a brief patient-reported measure of
453 shared decision making: Results from three clinical settings in the United States. *Health Expect*
454 2018;21:82-9.
- 455 28. Reeve E, Low LF, Shakib S, Hilmer SN. Development and Validation of the Revised Patients'
456 Attitudes Towards Deprescribing (rPATD) Questionnaire: Versions for Older Adults and Caregivers.
457 *Drugs Aging* 2016;33:913-28.
- 458 29. Reeve E, Low L-F, Hilmer SN. Beliefs and attitudes of older adults and carers about deprescribing
459 of medications: a qualitative focus group study. *The British journal of general practice : the journal of the*
460 *Royal College of General Practitioners* 2016;66:e552-60.
- 461 30. Reeve E, Low L-F, Hilmer SN. Attitudes of Older Adults and Caregivers in Australia toward
462 Deprescribing. *Journal of the American Geriatrics Society* 2019;67:1204-10.

- 463 31. Sun W, Tahsin F, Barakat-Haddad C, Turner JP, Haughian CR, Abbass-Dick J. Exploration of
464 home care nurse's experiences in deprescribing of medications: a qualitative descriptive study. *BMJ Open*
465 2019;9:e025606.
- 466 32. van Middelaar T, Ivens SD, van Peet PG, et al. Prescribing and deprescribing antihypertensive
467 medication in older people by Dutch general practitioners: a qualitative study. *BMJ open* 2018;8:e020871.
- 468 33. Zechmann S, Trueb C, Valeri F, Streit S, Senn O, Neuner-Jehle S. Barriers and enablers for
469 deprescribing among older, multimorbid patients with polypharmacy: an explorative study from
470 Switzerland. *BMC Fam Pract* 2019;20:64.
- 471 34. Frankel RM, Devers KJ. Study design in qualitative research--1: Developing questions and
472 assessing resource needs. *Educ Health (Abingdon)* 2000;13:251-61.
- 473 35. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering
474 implementation of health services research findings into practice: a consolidated framework for advancing
475 implementation science. *Implement Sci* 2009;4:50.
- 476 36. Reeve E, Shakib S, Hendrix I, Roberts MS, Wiese MD. The benefits and harms of deprescribing.
477 *Med J Aust* 2014;201:386-9.
- 478 37. Beuscart JB, Knol W, Cullinan S, et al. International core outcome set for clinical trials of
479 medication review in multi-morbid older patients with polypharmacy. *BMC Med* 2018;16:21.
- 480 38. Rankin A, Cadogan CA, In Ryan C, Clyne B, Smith SM, Hughes CM. Core Outcome Set for Trials
481 Aimed at Improving the Appropriateness of Polypharmacy in Older People in Primary Care. *J Am Geriatr*
482 *Soc* 2018;66:1206-12.
- 483 39. Eveleigh R, Grutters J, Muskens E, et al. Cost-utility analysis of a treatment advice to discontinue
484 inappropriate long-term antidepressant use in primary care. *Family practice* 2014;31:578-84.
- 485 40. Eveleigh R, Muskens E, Lucassen P, et al. Withdrawal of unnecessary antidepressant medication:
486 a randomised controlled trial in primary care. *BJGP open* 2018;1:bjgpopen17X101265.

487 41. Eveleigh R, Speckens A, van Weel C, Oude Voshaar R, Lucassen P. Patients' attitudes to
488 discontinuing not-indicated long-term antidepressant use: barriers and facilitators. *Therapeutic advances*
489 *in psychopharmacology* 2019;9:2045125319872344.

490 42. Gillespie P, Clyne B, Raymakers A, Fahey T, Hughes CM, Smith SM. REDUCING
491 POTENTIALLY INAPPROPRIATE PRESCRIBING FOR OLDER PEOPLE IN PRIMARY CARE:
492 COST-EFFECTIVENESS OF THE OPTI-SCRIPT INTERVENTION. *International journal of*
493 *technology assessment in health care* 2017;33:494-503.

494

495 **DESCRIPTIVE TITLE OF SUPPLEMENTAL MATERIAL**

496 Search strategy, list of articles and details on studies and measures

497

498

Table 1. Study characteristics (N=93)

Study characteristics	Number (%) of studies
Setting and patient characteristics	
Inpatient	19 (20)
Long-term care	27 (29)
Outpatient	51 (55)
Pharmacy	8 (9)
Other (emergency department, rehabilitative care, home care)	24 (26)
Older patients only	60 (65)
Methods	
Randomized study	21 (23)
Control group	42 (45)
Quantitative assessment	93 (100)
Qualitative assessment	18 (19)
Number of drug class(es) targeted by the interventions	
1	42 (45)
2	13 (14)
3	6 (6)
≥4	32 (34)
Classes of drugs targeted by the interventions	
Sedative-hypnotics	64 (69)
Antipsychotics	43 (46)
Antidepressants	36 (39)
Opioids	33 (36)
Anticholinergics	33 (36)
Proton pump inhibitors	35 (38)
Other drug class	35 (38)
Intervention type	
<i>Targeting patient</i>	44 (47)
Education	29 (31)
Drug substitution	8 (9)
Other	26 (28)
<i>Targeting provider</i>	85 (91)
Feedback / report card	9 (10)
Education	31 (33)
Guideline	20 (22)
Drug checklist	18 (19)
Drug review	40 (43)
Other clinical decision support	15 (16)
Pay for performance	1 (1)
Other	45 (48)

Total numbers for each characteristic are higher than the total number of studies, because some studies included more than one of these characteristics.

Table 2. Types and sources of measures

	Number (%) of measures	Number (%) of studies with ≥1 of the measure category / subcategory / source
Measure Type		
1. Appropriateness*	211 (51)	90 (97)
Cessation	171 (33)	79 (85)
Dose reduction	68 (13)	30 (32)
Switch for another drug	16 (3)	5 (5)
New prescription	14 (3)	3 (3)
Other	7 (1)	1 (1)
2. Utilization/ordering*	52 (10)	35 (38)
Cessation	16 (3)	10 (11)
Dose reduction	11 (4)	5 (5)
Switch for another drug	23 (5)	17 (18)
New prescription	21 (4)	13 (14)
Other	5 (1)	2 (2)
3. Intermediate outcome**	27 (5)	19 (20)
4. Outcome**	94 (18)	33 (35)
5. Patient-reported outcome	62 (12)	25 (27)
6. Patient-reported experience	15 (3)	12 (13)
7. Patient preferences	4 (1)	2 (2)
8. Provider-reported experience	16 (3)	11 (12)
9. Patient-provider interaction	4 (1)	4 (4)
10. Value (outcome/cost)	3 (1)	2 (2)
11. Cost	12 (2)	10 (11)
12. Other	11 (2)	10 (11)
Measure of unintended consequences	52 (10)	24 (26)
Definite unintended consequence	21 (4)	9 (10)
Possible unintended consequence	31 (6)	19 (20)
Measure source		
Patient-reported	117 (23)	33 (36)
Provider-reported	75 (15)	36 (39)
Medical / pharmacy record	349 (68)	86 (93)
Validated scale / questionnaire	66 (13)	25 (27)
Non-validated scale / questionnaire	30 (6)	16 (17)
Blinded assessment	92 (18)	11 (12)

*An appropriateness or utilization/ordering measure can be a combination of the subcategories, explaining that adding the subcategories results in more measures than the overall category.

**Not patient reported

Total number of measures: 511. Total number of unique studies: 93.

Table 3. Summary of measures used in the studies for each category and subcategory

1. Appropriateness (a), 2. utilization/ordering (b)
<i>Cessation:</i> a) number of patients with inappropriate drug ceased; b) mean number of prescriptions
<i>Dose reduction:</i> number of patients with: a) $\geq 50\%$ dose reduction of inappropriate drug; b) change in drug dose
<i>New prescription:</i> a) number of new inappropriate drugs; b) number of drugs restarted (appropriateness not assessed)
<i>Switch for another drug:</i> a) switches for alternative drug because of withdrawal; b) number with antidepressant as alternative
3. Intermediate outcome
Number of: deprescribing recommendations / drug alerts requiring an intervention
Proportion of: deprescribing recommendations accepted by patients / providers
Proportion of: patients with tapering plan developed / withdrawal attempt / receiving a deprescribing intervention
Reasons for: rejecting recommendation / not achieving deprescribing
4. Outcome
Healthcare services utilization (e.g., length of stay, hospitalization, outpatient visit)
Drug side effects / withdrawal signs (e.g., delirium, aggressive behavior, insomnia)
Adverse effects of drug cessation (e.g., hyperglycemia, fall, CVD event, seclusion room, physical restraints, death)
5. Patient-reported outcome
QoL / well-being / health status (EQ-5D-3L, 15D-HRQoL, Well-Being Questionnaire, 36-item Short Form Survey)
Functional status / activities of daily living (Groningen Activity Restriction Scale)
Withdrawal symptoms / drug side effects (SDS, BWSQ, Udvalg for Kliniske Undersogelser side effect rating scale)
Sleep quality / satisfaction (Pittsburgh Sleep Quality Index, Oviedo Sleep Questionnaire)
Gastrointestinal symptoms (Gastrointestinal Symptom Rating Scale, Gastroesophageal Reflux Disease Impact Scale)
Cognitive function (MoCA, MMSE, PAS-CIS; InterRAI-Long Term Care Facilities)
Psychopathology (Brief Symptoms Inventory, Hospital Anxiety and Depression Scale, Geriatric Depression Scale, CES-D)
Beliefs about drugs (Beliefs about Medicines Questionnaires) / Self-efficacy (Medication Reduction Self-efficacy Scale)
6. Patient-reported experience
Experience / satisfaction with the intervention (e.g., tapering process, implication in drug review, educational material)
Difficulties during the intervention / reasons for deprescribing failure (e.g., fears because of prior failed attempts, withdrawal)
7. Patient preferences
Proportion of patients who agreed / refused deprescribing; reason(s) for refusing
Preferences for the intervention
8. Provider-reported experience
Self-efficacy to describe / develop a deprescribing plan / implement a deprescribing plan
Satisfaction / experience / perception / difficulties / feasibility / acceptance / adoption / key messages of the intervention
Preferences for communication between providers (e.g., face-to-face, messages through electronic record)

Most useful part of the intervention (e.g., reminder message, tool, patient handout)
9. Patient-provider interaction
Personal interactions / discussions between patients and providers regarding deprescribing
Number of counseling occasions provided to each patient by the pharmacist / physician
Drug review with the patient
10. and 11. Cost-related
10. Value (outcome/cost): cost-utility (costs/QALYs) / cost-effectiveness (costs/number of potentially inappropriate drugs)
11. Costs: costs of: drugs / intervention (implementation, material (e.g., patient education brochure)) / healthcare services use
Unintended consequences
Switch for: substitute drug / additional drug / drug restarted for symptom control
Withdrawal signs or symptoms / worsening of symptoms treated by the deprescribed drug
Other adverse effects of deprescribing (e.g., hyperglycemia, CV events, QoL, death, fall)
Healthcare resource utilization (e.g., length of stay, hospitalization, outpatient visits)

Abbreviations: BWSQ, Benzodiazepine Withdrawal Symptom Questionnaire; CES-D, Centre for Epidemiological Studies Depression Scale; CV, cardiovascular; EQ-5D-3L, EuroQol five-dimensional three-level questionnaire; 15D-HRQoL, 15-dimensional health-related quality of life instrument; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; QALY, quality-adjusted life year; QoL, quality of life; PAS-CIS; Psychogeriatric Assessment Scales – Cognitive Impairment Scale; SDS, Severity of Dependence Scale.

Legend: Given that appropriateness and utilization/ordering measures are rather obvious and were ubiquitously used across studies, we only provide one example for each of their subcategories. For the other categories / subcategories, we synthesize all measures used across studies and provide examples of validated scales in brackets. Some measures are relevant for specific drugs only.

LEGENDS FOR FIGURES

Figure 1. Flow-chart of search result

