Validation of the LESS-CHRON criteria: reliability study of a tool for deprescribing in patients with multimorbidity

Aitana Rodríguez-Pérez,[1] Eva Rocío Alfaro-Lara,[1] María Isabel Sierra-Torres,[2] Ángela Villalba-Moreno,[1] María Dolores Nieto-Martin,[3] Mercedes Galván-Banqueri,[1] Bernardo Santos-Ramos[1]

[1]Unidad de Gestión Clínica de Farmacia, Hospital Universitario Virgen del Rocío, Seville, Spain
[2]Servicio de Farmacia, Área de Gestión Sanitaria Serranía de Málaga, Málaga, Spain
[3]Unidad de Gestión Clínica de Medicina Interna, Hospital Universitario Virgen del Rocío, Seville, Spain

Correspondence to
Mrs Aitana Rodríguez-Pérez, Unidad de Gestión Clínica de Farmacia, Hospital Universitario Virgen del Rocío, Seville 41013, Spain; aitanarp@gmail.com

EAHP Statement 4: Clinical Pharmacy Services.

ABSTRACT

Objective The 'LESS-CHRON criteria' (List of Evidence-Based Deprescribing for Chronic Patients criteria) is a newly created tool with 27 criteria to guide deprescribing. It was developed using a Delphi methodology. Each criterion consists of drugs and their indications, conditions under which deprescribing would be considered, a health variable to be monitored after deprescribing and a follow-up period. The aim of our study was to evaluate the reliability of the LESS-CHRON criteria in a population of patients with multimorbidity to determine the possible usefulness of this tool in clinical practice.

Methods We selected chronic patients with multimorbidity from an internal medicine unit who were older than 80 years old and were alive at the time of the study. To determine interobserver reliability, each professional (internist or hospital pharmacy specialist) applied the questionnaire under the same conditions and with the same resources. To determine intraobserver reliability, each health professional applied the tool at baseline and 2 months later. We measured interobserver and intraobserver reliability using the kappa coefficient. The proportion of overall agreement was also determined.

Results We obtained a moderate overall kappa (κ=0.46, 95% CI 0.36 to 0.55) for interobserver reliability, and good (κ=0.65, 95% CI 0.57 to 0.78) and moderate (κ=0.59, 95% CI 0.49 to 0.74) values for intraobserver reliability for the internist and pharmacist, respectively. The proportion of overall agreement was very high: 92% (range: 62%–100%) for the interobserver, and 94% (80%–100%) and 93% (63%–100%) for the internist and pharmacist, respectively.

Conclusions The LESS-CHRON criteria shows early promise as a reliable method to help guide deprescribing in patients with multimorbidity. Further, more complete testing with a larger sample of prescribers is needed.

INTRODUCTION

Due to the progressive ageing of the population and improvements in healthcare, patients with multimorbidity (PMM) are constantly increasing in hospital setting and primary care.[1–3] Multiple chronic medical conditions are associated with poor health outcomes: decreased quality of life,[4] longer hospital stay, higher cost of care and higher mortality.[5 6] Multimorbidity may also result in complex self-care needs; challenging organisational problems (accessibility, coordination and consultation time); increased use of emergency facilities; polypharmacy; and fragmented, costly and ineffective care.[7 8]

Available guidelines base pharmacotherapy recommendations on a single disease, and overlook the possibility of comorbidities and the consumption of other medications by the patient.[9] Treating patients with several diseases using a clinical practice guideline developed for each disease means using multiple drugs. This fact increases the chance of adverse drug reactions (ADR) and drug–drug interactions, which eventually add greater risks to the patient as a result of prescribing cascades and a deteriorating health.[10] Polypharmacy has also been associated with a lesser adherence to treatment, which contributes to greater healthcare expenditure.[11]

Different approaches to avoid the risks of polypharmacy in the elderly have been developed. Patient compliance, appropriateness of prescriptions and reconciliation process have been considered of vital importance for patients receiving a multiple drug treatment.[12] Deprescribing has been defined as a review and evaluation process of long-term therapeutic plans that aims to stop, substitute or reduce the dosage of those drugs that under certain clinical conditions can be considered unnecessary or with an unfavourable benefit to risk ratio.[13 14] It is a new issue to consider in the management of PMM, with the aim of reducing ADRs and drug burden and preventing prescribing cascades. However, strategies of deprescribing are currently in development and its application in clinical practice is, nowadays, limited.[15 16] Our group recently developed the LESS-CHRON criteria, a tool to guide the deprescribing process in PMM.[17] Every newly created tool should be assessed for its utility. Because there was no other deprescribing tool for PMM with which to compare the LESS-CHRON criteria with, it was not possible to assess its validity. Nevertheless, it was possible to assess its reliability. Reliability provides information on the reproducibility of the results obtained by the procedure used for measurement; it is the degree of stability achieved in the results when a measurement is repeated under identical conditions. The utility of a new tool should be check by the professionals who are going to use it in the clinical practice.[18] As previously done for other similar tools (STOPP/START (Screening Tool of Older Person's

Prescriptions/Screening Tool to Alert doctors to Right Treatment),[19] OncPal deprescribing guideline[20]), a reliability study to validate the new criteria was designed.

The aim of this study was to evaluate the reliability of the LESS-CHRON criteria in a population of PMM to determine the possible usefulness of this tool in normal clinical practice.

METHODS
LESS-CHRON criteria
The 'LESS-CHRON criteria' is a newly created tool to guide deprescribing in PMM. It was developed through a literature review, followed by a Delphi methodology.[17] It contains a list of 27 criteria. Each criterion consists of drugs used in a specific clinical setting, a deprescribing condition, a health variable to monitor after deprescribing and a follow-up period.

Study design
An interobserver and intraobserver reliability study was performed. The observers were a specialist in a hospital pharmacy and a specialist in internal medicine (internist). They were selected for their recognised professional experience in chronic patient care and for their expertise in deprescribing. They both worked in the hospital where the study was conducted, so they knew how to use the clinical database. They were also associated with the PMM investigative research group of the hospital. The internist was an academic doctor with 8 years of clinical experience who used to treat outpatients as well as inpatients with multimorbidity. The pharmacist was a researcher with 4 years of clinical experience who was responsible for clinical validation of treatments in the internal medicine unit of the hospital.

They were provided with a LESS-CHRON worksheet, which contained only the first two parameters of each criterion, as monitoring and follow-up were not going to be evaluated in the study (online supplementary material 1). Moreover, resources on the value of the different scales and indexes that appear in the tool (Barthel, Pfeiffer and Profund[21–23]) were also provided.

With the objective of determining interobserver reliability, each healthcare professional applied the tool under the same conditions and using the same resources. The necessary information from the patient was obtained from computerised clinical histories. The healthcare professionals had to indicate for each patient whether the criterion was fulfilled (scoring it as '1') or not (scoring it as '0'). It means that they had to check if a target drug had been prescribed for the indication collected in the tool and if the clinical conditions of the patient are those that the tool defined as 'deprescribing conditions'. If both were present in a patient, they considered the criterion as fulfilled. This was the main variable of the study.

To establish intraobserver reliability, each healthcare professional reapplied the tool in the same sample of patients and with the same records and information 2 months later.

Patients
PMMs who attend a tertiary hospital internal service unit either as inpatients or outpatients were selected. They were identified with the support of the Clinical Documentation Service of the hospital by a smart search in the digital clinical history database.

The inclusion criteria were 80 years old or older, have a summary discharge from the internal medicine unit in September 2014–May 2015, presenting with one of the pathologies considered a clinical indication for the drugs included in the LESS-CHRON criteria and have active prescriptions in their health records.

Patients whose pathologies were diagnosed after the initial admission or who were no longer alive at the time of the study were excluded.

Statistical analysis and secondary outcomes
Interobserver and intraobserver reliability was determined using the kappa coefficient (κ), a measure of the extent of agreement beyond what would be expected by chance alone. The kappa coefficient was interpreted as slight if 0.2 or less, fair if 0.21–0.4, moderate if 0.41–0.6, substantial if 0.61–0.8 and good if 0.81–1.[24]

The minimum number of patients required for the inter-reliability study was calculated from a general assumption of the number of subjects required in a two-rater study to detect a statistically significant κ on a dichotomous variable.[24] We assumed the null hypothesis value of kappa to be 0.00 and aimed a 90% power to detect a kappa value of at least 0.6, estimating the internist and pharmacist would have greater than 90% concordance. This calculation resulted in at least 30 patients. Taking into account this value and the number of patients previously used in similar studies,[20 25–27] a sample of 50 patients was estimated.

Table 1 Interobserver agreement: internist (evaluator 1)/ hospital pharmacy specialist (evaluator 2) (n=50)

Criterion	α	β	γ	δ	ppos	pneg	κ	PO
A1	1	0	1	48	0.67	0.99	0.66	0.98
A2	0	0	0	50	NA	1.00	NA	1.00
A3	1	1	1	47	0.50	0.98	0.48	0.96
A4	2	1	4	43	0.44	0.95	0.40	0.90
B1	2	6	0	42	0.40	0.93	0.36	0.88
B2	8	5	3	34	0.67	0.89	0.56	0.84
B3	1	6	0	43	0.25	0.93	0.22	0.88
B4	1	2	0	47	0.50	0.98	0.49	0.96
C1	7	1	18	24	0.42	0.72	0.24	0.62
C2	1	5	6	38	0.15	0.87	0.03	0.78
C3	2	1	2	45	0.57	0.97	0.54	0.94
C4	0	1	0	49	0.00	0.99	0.00*	0.98
D1	0	0	0	50	NA	1.00	NA	1.00
D2	0	0	0	50	NA	1.00	NA	1.00
D3	4	0	7	39	0.53	0.92	0.47	0.86
D4	1	0	2	47	0.50	0.98	0.49	0.96
E1	0	0	0	50	NA	1.00	NA	1.00
E2	1	2	1	46	0.40	0.97	0.37	0.94
F1	0	6	0	44	0.00	0.94	0.00*	0.88
F2	10	6	4	30	0.67	0.86	0.52	0.80
F3	10	3	5	32	0.71	0.89	0.60	0.84
F4	2	3	4	41	0.36	0.92	0.29	0.86
F5	0	0	1	49	0.00	0.99	0.00*	0.98
F6	0	0	0	50	NA	1.00	NA	1.00
F7	1	0	0	49	1.00	1.00	1	1.00
F8	0	0	0	50	NA	1.00	NA	1.00
G1	0	2	0	48	0.00	0.98	0.00*	0.96
Global	55	53	59	1183	0.50	0.96	0.46	0.92

α: Both evaluators scored the criterion as being fulfilled.
β: Evaluator 1 scored the criterion as being fulfilled and evaluator 2 scored the criterion as not fulfilled.
γ: Evaluator 2 scored the criterion as being fulfilled and evaluator 1 scored the criterion as not fulfilled.
δ: Both evaluators agreed the criterion was not fulfilled.
*Inadequate variability between groups.
NA, not applicable; pneg, proportion of negative agreement; PO, global per cent agreement; ppos, proportion of positive agreement.

Table 2 Interobserver agreement for each group of drugs of the tool

Group of drugs	n	α	β	γ	δ	ppos	pneg	κ (95% CI)	PO
A. Alimentary tract and metabolism	200	4	2	6	188	0.50	0.98	0.48 (0.13 to 0.83)	0.96
B. Blood and blood forming organs	200	12	19	3	166	0.52	0.94	0.47 (0.26 to 0.68)	0.89
C. Cardiovascular system	200	10	8	26	156	0.37	0.90	0.28 (0.07 to 0.50)	0.83
D. Genitourinary system	200	5	0	9	186	0.53	0.98	0.51 (0.19 to 0.82)	0.955
E. Musculoskeletal system	100	1	2	1	96	0.40	0.98	0.39 (−0.30 to 1.07)	0.97
F. Nervous system	400	23	18	14	345	0.59	0.96	0.55 (0.39 to 0.70)	0.92
G. Respiratory system	50	0	2	0	48	0.00	0.98	0	0.96

n: Number of total evaluations for each group.
α: Both evaluators scored the criteria for this group as being fulfilled.
β: Evaluator 1 scored the criteria for this group as being fulfilled and evaluator 2 scored the criteria as not fulfilled.
γ: Evaluator 2 scored the criteria for this group as being fulfilled and evaluator 1 scored the criteria as not fulfilled.
δ: Both evaluators agreed the criteria for this group were not fulfilled.
*Inadequate variability between groups.
pneg, proportion of negative agreement; PO, global per cent agreement; ppos, proportion of positive agreement.

However, the sample size used to determine intraobserver reliability was 15 patients, which is the proportion from the sample used by similar studies.[25]

The proportion of positive agreement (ppos), interpreted as the likelihood that the criterion was fulfilled, and the proportion of negative agreement (pneg), interpreted as the likelihood that the criterion was not fulfilled, were calculated. Moreover, the global per cent agreement was calculated.[28] Additionally, the grade of agreement in the inter-reliability study for each group of drugs contained in the tool was analysed.

Microsoft Office Excel V.2010 and IBM SPSS Statistics V.21 were used for all statistical analyses.

RESULTS
Interobserver reliability
After applying the inclusion criteria, 623 patients were identified by the smart search. Every patient had at least one drug included in the tool. Fifty patients were randomly selected for evaluation. There were 29 women and 21 men, with an average age of 85 years old. The average (±SD) number of drugs included in patients' treatments was 11 (±4), and 5 (±2) of these drugs were included in the tool.

To determine interobserver reliability, data were classified, processed and tabulated as shown in table 1.

The kappa coefficient values varied widely between the different criteria, with a range of 0–1. The global kappa value was 0.46 (95% CI 0.36 to 0.55), which means that interobserver reliability was moderate. Six criteria were considered not applicable to any patient. However, the global per cent agreement value was 92% (range of values for the different items: 62%–100%). The global ppos was 0.5 (range: 0–0.71) and the global pneg was 0.96 (range: 0.72–1).

On agreement between observers in each group of drugs, the kappa coefficient value was fair in the criteria that contain drugs for the cardiovascular system (κ=0.28, 95% CI 0.07 to 0.50) and the musculoskeletal system (κ=0.39, 95% CI −0.30 to 1.07). However, the strength of agreement was moderate in the criteria for the alimentary tract and metabolism (κ=0.48, 95% CI 0.13 to 0.83), blood and blood forming organs (κ=0.47, 95% CI 0.26 to 0.68), genitourinary system (κ=0.51, 95% CI 0.19 to 0.82) and nervous system (κ=0.55, 95% CI 0.39 to 0.70). The criterion that contains drugs for the respiratory system did not have enough variability in agreement between groups to result in a valid value of kappa.

The global per cent agreement value in each group varied from 83% to 97% (table 2).

Intraobserver reliability
Fifteen patients out of the 50 patients included in the study were randomly selected. Tables 3A and 3B summarise the results.

In the case of the internist, the global kappa value for the intraobserver study was substantial (κ=0.65, 95% CI 0.57 to 0.78). However, the global per cent agreement reached 94% (range: 80%–100%). The global ppos was 0.68 (range: 0–1) and the global pneg was 0.97 (range: 0.82–1).

In the case of the pharmacist, the intraobserver agreement was moderate (κ=0.59, 95% CI 0.49 to 0.74). The global per cent agreement was 93% (range: 63%–100%). The global ppos was 0.63 (range: 0–1) and the global pneg was 0.96 (range: 0.74–1).

DISCUSSION
The present study concludes that the LESS-CHRON criteria shows early promise as a reliable method to help guide deprescribing in PMM. To our knowledge, the 'LESS-CHRON criteria' is the first evidence-based tool for deprescribing in PMM that has been developed and further validated. The methodology of this work is well established. There are similar studies in which reliability has been measured, most of them using the 'STOPP criteria'[19 26 27] and one using the 'OncPal deprescribing guideline'.[20] The global kappa which resulted in interobserver reliability in our study (κ=0.46) fell below the values reached in those studies. The three studies which measured the reliability of the 'STOPP criteria' reached kappa values of 0.7–1. It could be because the patients were consciously selected in those studies to present a high incidence of potentially inappropriate medicines, as authors explained,[26 27] while in the present study patients were randomly selected from a population who were treated in the internal medicine service. This means that it was more difficult to detect conditions that allowed drugs to be deprescribed, but that it provided more external validity. Unlike the other studies, in our study there was no standard reference to compare the results with after applying the LESS-CHRON criteria.[25] In those studies, the results of the observers were compared with those obtained by an expert panel,[20] academic researchers[26] or the creator of the criteria[27] when they applied the criteria to the same patients. Thus, the kappa value in our study is the concordance between two clinicians after applying the tool, but does not mean that the tool is poorly applied and that it is similar to what occurred in the first validation study of the STOPP criteria.[19] The kappa index has an important limitation: the lower the prevalence of the phenomenon of interest, the lower the kappa value.[28] In our study many criteria were not considered by the observers to be fulfilled because there were

Original article

Table 3A Intraobserver agreement for the internist (n=15)

Criterion	α	β	γ	δ	ppos	pneg	κ	PO
A1	1	0	0	14	1.00	1.00	1.00	1.00
A2	0	0	0	15	NA	1.00	NA	1.00
A3	0	0	1	14	0.00	0.97	0.00*	0.93
A4	0	1	0	14	0.00	0.97	0.00*	0.93
B1	3	0	0	12	1.00	1.00	1.00	1.00
B2	3	1	0	11	0.86	0.96	0.81	0.93
B3	2	0	1	12	0.80	0.96	0.76	0.93
B4	1	1	1	12	0.50	0.92	0.42	0.87
C1	2	2	1	10	0.57	0.87	0.44	0.80
C2	1	2	0	12	0.50	0.92	0.44	0.87
C3	1	0	1	13	0.67	0.96	0.63	0.93
C4	0	0	0	15	NA	1.00	NA	1.00
D1	0	0	0	15	NA	1.00	NA	1.00
D2	0	0	0	15	NA	1.00	NA	1.00
D3	0	1	0	14	0.00	0.97	0.00*	0.93
D4	0	0	0	15	NA	1.00	NA	1.00
E1	0	0	1	14	0.00	0.97	0.00*	0.93
E2	0	1	0	14	0.00	0.97	0.00*	0.93
F1	0	1	1	13	0.00	0.93	0.00*	0.87
F2	5	1	2	7	0.77	0.82	0.59	0.80
F3	4	1	2	8	0.73	0.84	0.57	0.80
F4	4	0	1	10	0.89	0.95	0.84	0.93
F5	0	0	0	15	NA	1.00	NA	1.00
F6	0	0	0	15	NA	1.00	NA	1.00
F7	0	0	1	14	0.00	0.97	0.00*	0.93
F8	0	0	0	15	NA	1.00	NA	1.00
G1	0	0	0	15	NA	1.00	NA	1.00
Global	27	12	13	353	0.68	0.97	0.65	0.94

α: Pre and post, the criterion was considered fulfilled.
β: Pre fulfilled and post not fulfilled.
γ: Post fulfilled and pre not fulfilled.
δ: Pre and post not fulfilled.
*Inadequate variability between groups.
NA, not available; pneg, proportion of negative agreement; PO, global per cent agreement; ppos, proportion of positive agreement.

Table 3B Intraobserver agreement for the pharmacist (n=15)

Criterion	α	β	γ	δ	ppos	pneg	κ	PO
A1	1	0	0	14	1.00	1.00	1.00	1.00
A2	0	0	0	15	NA	1.00	NA	1.00
A3	0	0	0	15	NA	1.00	NA	1.00
A4	1	2	0	12	0.50	0.92	0.44	0.87
B1	1	0	0	14	1.00	1.00	1.00	1.00
B2	2	3	0	10	0.57	0.87	0.47	0.80
B3	0	1	0	14	0.00	0.97	0.00*	0.93
B4	1	0	0	14	1.00	1.00	1.00	1.00
C1	3	4	0	8	0.60	0.80	0.44	0.73
C2	0	1	1	13	0.00	0.93	0.00*	0.87
C3	0	3	0	12	0.00	0.89	0.00*	0.80
C4	0	0	0	15	NA	1.00	NA	1.00
D1	0	0	0	15	NA	1.00	NA	1.00
D2	0	0	0	15	NA	1.00	NA	1.00
D3	3	0	1	11	0.86	0.96	0.81	0.93
D4	1	0	0	14	1.00	1.00	1.00	1.00
E1	0	0	0	15	NA	1.00	NA	1.00
E2	0	1	0	14	0.00	0.97	0.00*	0.93
F1	0	0	0	15	NA	1.00	NA	1.00
F2	3	2	3	7	0.55	0.74	0.29	0.67
F3	4	2	1	8	0.73	0.84	0.57	0.80
F4	3	0	2	10	0.75	0.91	0.67	0.87
F5	0	0	0	15	NA	1.00	NA	1.00
F6	0	0	0	15	NA	1.00	NA	1.00
F7	0	0	0	15	NA	1.00	NA	1.00
F8	0	0	0	15	NA	1.00	NA	1.00
G1	0	0	0	15	NA	1.00	NA	1.00
Global	23	19	8	355	0.63	0.96	0.59	0.93

α: Pre and post, the criterion was considered fulfilled.
β: Pre fulfilled and post not fulfilled.
γ: Post fulfilled and pre not fulfilled.
δ: Pre and post not fulfilled
*Inadequate variability between groups.
NA, not applicable; pneg, proportion of negative agreement; PO, global per cent agreement; ppos, proportion of positive agreement.

no patients with those characteristics. Conversely, the agreement was almost 100% in considering them not to be fulfilled, as can be shown in table 1 (high pneg values). Also, the health professionals who applied the LESS-CHRON criteria in this study had different background, skills and aptitudes. Their different profiles provided more strength to the method, although this may have contributed to a kappa value lower than in other studies.

In the case of the intraobserver reliability, the results were almost similar between the two professionals (tables 3A and 3B).

With respect to the criteria which were not fulfilled by any patients, two of them were because the drug was not active in any treatment (A2: acarbose, F8: citicoline). The other four were because no patient fulfilled the condition of deprescribing (D1, D2: no patients with anticholinergics for urinary incontinence who were users of continence pads or were under anticholinesterase treatment; E1: no patients with more than 5 years of bisphosphonate treatment for primary prevention; and F6: no patients with anticholinesterases and memantine for Alzheimer disease).

There are some limitations to the method used. A larger sample would have allowed more available conditions to verify all the criteria. Another methodological limitation is the fact that

there were only two observers from different professions and with different years of experience in the management of PMM. It might have been better if prior to the study they had participated in a training session driven by the researchers to apply the tool.[25]

The application of LESS-CHRON using only clinical histories is another limitation. There are several criteria which need to utilise a clinical scale or to check some clinical judgements that are not always well recorded in clinical histories. As an example, most of the criteria used for the central nervous system drugs probably need a clinical assessment of patients to check behavioural stability, mood recovery or grade of anxiety or insomnia. This might be the explanation for the low agreement reached in this group of criteria (F1–F4).

There are criteria which need to be clarified according to the differences observed in the application of the tool, and which are responsible for the fair kappa values obtained in the criteria that contained drugs for the cardiovascular system and musculoskeletal system. For instance, it would be better if the tool included a brief explanation about the first-line treatment of antihypertensives (C1); what is considered primary and secondary prevention in the use of statins and bisphosphonates (C2, E2); and what age is considered as risk factor in the use of acetylsalicylic acid (B3).

The research group analysed and studied them, and the results introduced some clarifications in those criteria which need to be improved in the future.

Although time constraints often limit the use of this kind of tool in everyday clinical practice, the possibility of incorporating the LESS-CHRON criteria into a computer-aided decision support software system is currently being investigated. Aware of the importance of clinically validating this new deprescribing tool, a study in collaboration between the Spanish Society of Internal Medicine and the Spanish Society of Hospital Pharmacy is in progress to develop a deprescribing clinical trial using the LESS-CHRON criteria. The intervention group will be a group of internists from different Spanish hospitals who will apply the LESS-CHRON criteria. The control group will apply the usual clinical practice of revising the therapeutic plans of patients. The number of drugs deprescribed, adherence, cognitive status and quality of life will be compared between each other. The study will also allow validation of the other two parameters of each LESS-CHRON criterion (monitoring and follow-up).

CONCLUSIONS

The LESS-CHRON criteria shows early promise as a reliable method to help guide deprescribing in PMM. Further, more complete testing with a larger sample of prescribers is needed.

What this paper adds

What is already known on this subject
► Deprescribing is a patient-centred process that can benefit from a drugs list criteria to support and guide the process.
► The LESS-CHRON criteria (List of Evidence-Based Deprescribing for Chronic Patients criteria) is a tool created to guide deprescribing in patients with multimorbidity.

What this study adds
► The LESS-CHRON criteria shows early promise as a reliable method to help guide deprescribing in patients with multimorbidity.

REFERENCES

1 Marengoni A, Angleman S, Melis R, et al. Aging with multimorbidity: a systematic review of the literature. *Ageing Res Rev* 2011;10:430–9.
2 Fortin M, Bravo G, Hudon C, et al. Prevalence of multimorbidity among adults seen in family practice. *Ann Fam Med* 2005;3:223–8.
3 Bernabeu-Wittel M, Jadad A, Moreno-Gaviño L, et al. Peeking through the cracks: an assessment of the prevalence, clinical characteristics and health-related quality of life (HRQoL) of people with polypathology in a hospital setting. *Arch Gerontol Geriatr* 2010;51:185–91.
4 Fortin M, Bravo G, Hudon C, et al. Relationship between multimorbidity and health-related quality of life of patients in primary care. *Qual Life Res* 2006;15:83–91.
5 Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med* 2007;22(Suppl 3):391–5.
6 Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med* 2002;162:2269–76.
7 Bayliss EA, Steiner JF, Fernald DH, et al. Descriptions of barriers to self-care by persons with comorbid chronic diseases. *Ann Fam Med* 2003;1:15–21.
8 Fortin M, Soubhi H, Hudon C, et al. Multimorbidity's many challenges. *BMJ* 2007;334:1016–7.
9 Mortazavi SS, Shati M, Keshtkar A, et al. Defining polypharmacy in the elderly: a systematic review protocol. *BMJ Open* 2016;6:e010989.
10 Boyd CM, Darer J, Boult C, et al. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005;294:716–24.
11 Maher RL, Hanlon J, Hajjar ER. Clinical consequences of polypharmacy in elderly. *Expert Opin Drug Saf* 2014;13:57–65.
12 Alfaro Lara ER, Vega Coca MD, Galván Banqueri M, et al. Selection of tools for reconciliation, compliance and appropriateness of treatment in patients with multiple chronic conditions. *Eur J Intern Med* 2012;23:506–12.
13 Woodward MC. Deprescribing: Achieving Better Health Outcomes for Older People through Reducing Medications. *Journal of Pharmacy Practice and Research* 2003;33:323–8.
14 Rodríguez Pérez A, Alfaro Lara ER, Nieto Martín MD, et al. Deprescribing in patients with multimorbidity: a necessary process. *Eur J Intern Med* 2015;26:e18–19.
15 Scott IA, Hilmer SN, Reeve E, et al. Reducing inappropriate polypharmacy: the process of deprescribing. *JAMA Intern Med* 2015;175:827–34.
16 Scott IA, Le Couteur DG. Physicians need to take the lead in deprescribing. *Intern Med J* 2015;45:352–6.
17 Rodríguez-Pérez A, Alfaro-Lara ER, Albiñana-Perez S, et al. Novel tool for deprescribing in chronic patients with multimorbidity: List of Evidence-Based Deprescribing for Chronic Patients criteria. *Geriatr Gerontol Int* 2017;17:2200–7.
18 Sánchez Fernández P, Armas A I, Fuentelsaz Gallego C, et al. [Reliability of measurement instruments in the health sciences]. *Enferm Clin* 2005;15:227–36.
19 Gallagher P, Ryan C, Byrne S, et al. STOPP (Screening Tool of Older Person's Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment). Consensus validation. *Int J Clin Pharmacol Ther* 2008;46:72–83.
20 Lindsay J, Dooley M, Martin J, et al. The development and evaluation of an oncological palliative care deprescribing guideline: the 'OncPal deprescribing guideline'. *Support Care Cancer* 2015;23:71–8.
21 Mahoney FI, Barthel DW. FUNCTIONAL EVALUATION: THE BARTHEL INDEX. *Md State Med J* 1965;14:61–5.
22 Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *J Am Geriatr Soc* 1975;23:433–41.
23 Bernabeu-Wittel M, Ollero-Baturone M, Moreno-Gaviño L, et al. Development of a new predictive model for polypathological patients. The PROFUND index. *Eur J Intern Med* 2011;22:311–7.
24 Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85:257–68.
25 Galván-Banqueri M, de la Higuera-Vila L, Vega-Coca MD, et al. Reliability of a questionnaire for pharmacological treatment appropriateness in patients with multiple chronic conditions. *Eur J Intern Med* 2013;24:420–4.
26 Ryan C, O'Mahony D, Byrne S. Application of STOPP and START criteria: interrater reliability among pharmacists. *Ann Pharmacother* 2009;43:1239–44.
27 Gallagher P, Baeyens JP, Topinkova E, et al. Inter-rater reliability of STOPP (Screening Tool of Older Persons' Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment) criteria amongst physicians in six European countries. *Age Ageing* 2009;38:603–6.
28 Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551–8.