

# Consensus Validation of the FORTA (Fit FOR The Aged) List: A Clinical Tool for Increasing the Appropriateness of Pharmacotherapy in the Elderly

Alexandra M. Kuhn-Thiel · Christel Weiß · Martin Wehling ·  
The FORTA authors/expert panel members

Published online: 19 December 2013

© The Author(s) 2014. This article is published with open access at Springerlink.com

## Abstract

**Background** Multimorbidity and polypharmacy represent a major problem for elderly patients; improvement of medication schemes is important and listing approaches (e.g. Beers list) are considered to be potentially useful.

**Objectives** The aim of this study was to perform expert consensus validation of the FORTA (Fit FOR The Aged) List, a drug classification combining positive and negative labelling of drugs chronically prescribed to elderly patients.

**Methods** A two-round Delphi procedure was conducted involving 20 experts, 17 geriatric internists and 3 geriatric psychiatrists from Germany and Austria, evaluating the labels assigned to 190 substances or substance groups. These labels ranged from A (indispensable), B (beneficial), C (questionable) to D (avoid), depending on the state of evidence for safety, efficacy and overall age-appropriateness. The experts were also requested to suggest additional substances and indication areas for assessment and possible inclusion in the FORTA List. A weighted (corrected)

consensus coefficient was generated for each substance to reflect (1) agreement with the original label, and (2) distribution among raters' labels.

**Results** The overall consensus for all items and raters was 92 % (corrected). For 54/190 items (28.4 %), a unanimous response was achieved as to the original author-based FORTA label choice. Twenty-four substances (12.6 %) fell short of the consensus cutoff and were re-evaluated in a second round. This yielded confirmation of 171/190, or 90 %, of the original author-based FORTA labels. A total of 35 new substances were also accepted for the FORTA List. Drugs used for dementia and dementia syndromes provoked particular response heterogeneity.

**Conclusion** The FORTA List now reflects a wider consensus among experts, increasing its validity for clinical use. It represents a tool to improve the quality of drug prescription in older patients by identifying both inappropriate and omitted drugs, and thus overtreatment and undertreatment. The validation of FORTA's impact on clinical endpoints has yielded promising preliminary results, to be corroborated in ongoing larger trials.

The FORTA authors/expert panel members are listed in the [Appendix](#).

**Electronic supplementary material** The online version of this article (doi:10.1007/s40266-013-0146-0) contains supplementary material, which is available to authorized users.

A. M. Kuhn-Thiel · M. Wehling (✉)  
Institute for Experimental and Clinical Pharmacology and Toxicology, Department of Clinical Pharmacology, Center for Geriatric Pharmacology, Medical Faculty Mannheim, University of Heidelberg, Maybachstr. 14, 68169 Mannheim, Germany  
e-mail: martin.wehling@medma.uni-heidelberg.de

C. Weiß  
Department of Medical Statistics, Biomathematics and Information Processing, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

## 1 Introduction

The increase in life expectancy is triggering dramatic demographic changes in industrialized countries. Figures for 2010 provided by the US Census Bureau [1] indicate that the population aged 65 years and older has increased over the past decade from 35.0 million in 2000 to 40.3 million in 2010, representing 13.0 % of the total population. In Germany, this percentage is at 20 % for 2009 [2].

Elderly multimorbid patients are more likely to receive multiple drug treatments (polypharmacy) compared with younger patients. It has been demonstrated that patients

aged 65 years and older take five or more drugs in 44 % (male) and 57 % (female) of cases, and ten or more drugs in 12 % of cases [3]. Each medical guideline recommends an average of three medications. According to figures from 1998, persons over 80 years of age have an average of three diagnoses; this means  $3 \times 3 = 9$  medications in elderly patients, which is also reflected in real life [4].

Multimorbidity and polypharmacy often harbour unpredictable dangers due to age-related alterations in pharmacokinetics and pharmacodynamics [5], adverse drug reactions [6], which may trigger the so-called prescribing cascade [7], drug–drug or drug–disease interactions, problems with dosages, medication errors, and even death [8]. This translates, for example, into approximately 2.1 million side effect-related hospital admissions and 100,000 deaths per year in the US, out of a total population of 265 million inhabitants, and costs of approximately 1.5–4 billion dollars per year [9]. The risk of potential drug interactions necessarily increases with the number of drugs prescribed [10].

Geriatric medicine is a rapidly growing discipline in the Western world [11, 12]. The paucity of evidence-based guidelines and clinical studies for the elderly is alarming and contributes to the challenges of rationalistic drug prescribing in elderly patients. Many elderly persons display limitations in their physical and mental capacities [13], nearly always precluding their inclusion in clinical drug trials. The sometimes rigid adherence to available guidelines presents another significant problem as there are virtually no evidence-based guidelines for this very heterogeneous group; the underlying assumption ‘one guideline fits all’ simply does not work in this age bracket [14]. These limitations necessitate the development of criteria or concepts for safer and more efficient drug use in the elderly, ideally amalgamated at an international level.

In response to these challenges, many countries have begun to develop strategies for the safer prescribing of medications in elderly patients [15]. A precedent was set by the Beers Criteria [16] which classify potentially inappropriate medications (PIMs) according to three categories describing the degree of inappropriateness; this listing has undergone several updates, recently winning the support of the American Geriatric Society [17]. The German counterpart, the PRISCUS (PIM) List, was published in 2010 [18]. Such ‘negative lists’ have proven to be quite practicable, but still lack confirmation as to effectiveness at the clinical endpoint level [19].

Gallagher et al. introduced the STOPP (Screening Tool of Older Persons’ Prescriptions)/START (Screening Tool to Alert Doctors to the Right Treatment) Criteria in 2008 [20]. The STOPP criteria allow the detection of potential overtreatment and place a special emphasis on drug–

disease interactions; the START criteria serve to assist the physician in targeting potential errors of omission by pinpointing treatments that may be indicated but not prescribed [21]. It has been suggested in the literature that the STOPP criteria may have a higher sensitivity than the Beers criteria for detecting PIMs [22].

The FORTA (Fit fOR The Aged) classification system was proposed in 2008 [23, 24] as a tool for aiding physicians from participating countries (initially Germany) in screening for unnecessary, inappropriate or harmful medications and drug omissions in older patients in an everyday clinical setting. It is the first classification system in which both negative and positive labelling are combined at the level of individual drugs or drug groups. As it aims at the individual indications (implicit listing requiring patient characteristics/diagnoses) it is thus clearly different from negative lists which focus on major problems in drug prescribing (errors of commission or omission, or risky medications within all frequently used drug classes) that should rather be avoided when prescribing to geriatric patients because of age-related changes (explicit lists largely independent of individual patient characteristics). FORTA is evidence-based and real-life oriented. Factors such as adherence issues, age-dependent tolerance and frequency of relative contraindications are given due consideration since strict and citable evidence, as typically derived from randomized clinical trials, is still rare for this population, although important if available. A medication can receive different FORTA labels for different indications (indication-dependent). FORTA does not take the place of individual therapeutic considerations or decisions. Contraindications always take precedence over the FORTA-classification. The system does allow for exceptions.

The FORTA classes are defined as follows:

- *Class A* (A-bsolutely) = indispensable drug, clear-cut benefit in terms of efficacy/safety ratio proven in elderly patients for a given indication
- *Class B* (B-eneficial) = drugs with proven or obvious efficacy in the elderly, but limited extent of effect or safety concerns
- *Class C* (C-areful) = drugs with questionable efficacy/safety profiles in the elderly, to be avoided or omitted in the presence of too many drugs, lack of benefits or emerging side effects; review/find alternatives
- *Class D* (D-on’t) = avoid in the elderly, omit first, review/find alternatives

The FORTA List is a compilation of 190 medications (primarily long-term treatment, exceptions are noted) [25] most frequently prescribed in older patients, aligned to 20 main indication groups. Each substance or group is assigned a FORTA class A, B, C or D. In cases in which homogeneity

was considered to be high and of lesser significance with respect to other aspects of the group of drugs concerning age-related issues, similar drugs were grouped and assessed as such (angiotensin-converting enzyme [ACE] inhibitors as an example). If individual compounds were considered to be heterogenous, rating was performed for individual drugs (acetylsalicylic acid, clopidogrel—although both are considered as platelet inhibitors). This original version of the FORTA list is part of reference [25] and was created by the book authors, including the original author of the method. The rating was opinion-based by an integrative approach which comprised available study evidence as cited in the book. It thus was also evidence-based where such evidence was available. Although exact-use data are as yet missing, the original FORTA list is increasingly recognized in Germany, as suggested by the fact that the 3rd edition of the seminal book (2013) had to be published within 3 years after the first (2010); for this survey, the 2nd edition (2011) was relevant [25].

Our aims, within the context of a two-round Delphi Consensus procedure, included the rater-based confirmation/determination of labels for 190 items in the original author-based FORTA List, and identification and labeling of new indications/substances. As a consensus-based approach this process reflects the rating of many experts, drawing on both available evidence and personal experience/opinion.

The consensus validation of the FORTA List represents the first phase of a two-part development programme funded by the German Research Foundation; a clinical study is running to test the impact of the FORTA system on the quality of pharmacotherapy and clinical endpoints in 400 patients from two German geriatric clinics by implementing the FORTA List.

## 2 Methods

The consensus validation procedure included

1. Review of the available literature and examples of practical applications of the Delphi method.
2. Recruitment of experts in the German-speaking countries (Germany, Austria and Switzerland) representing geriatric internists/geriatricians and geriatric psychiatrists with extensive clinical experience in the pharmacotherapy of (multimorbid) elderly patients; high academic status; prominent standing in the leading geriatric/psychiatric medical associations; and number, quality and relevance of experts' publications. The selection was based on available information on the Internet in an iterative, semi-quantitative process by two of the authors (AKT and MW).

3. Round 1: The FORTA List was adapted from its original publication form [25] to a questionnaire and sent to the experts via e-mail (see original survey as Electronic Supplementary Material [ESM] 2). Participants were requested to review the instructions on how to apply the FORTA principle; study the author-based labels (A–D) for each item and provide their own FORTA labels (or abstention) whenever in disagreement; and make comments. In a separate section, the experts were requested to suggest new substances/indications with labels, to augment the FORTA List. All experts had been exposed to the related book [25] as the common, although certainly not exclusive, base of evidence compiled from the literature.
4. Statistical analysis based on Round 1 input was performed as follows: the Likert scale is often favoured for consensus procedures, as well as means, median and mode ('central tendency' indicators) [26, 27]. For evaluating the FORTA labelling system, we adapted the Likert scale with the aim of devising an algorithm combining collective/central tendency regarding the original labels with impact of distribution/dispersion of raters' labels. To achieve this, the percentage of raters' labels (excluding abstentions) agreeing with the original author-based labels was calculated, both overall and for each item separately. The resulting percentages were then weighted to generate a corrected consensus coefficient (cons\_corr, definition see ESM 1, FORTA list, p. 41) for each item reflecting the degrees of deviation between the experts' individual FORTA ratings. Although at first glance seemingly arbitrary, those weighing factors appear, for our purposes, plausible [28–30]. This ultimately allows the actual assignment of FORTA class values to the substances in question. The weighting system, reflecting degrees of deviation, was expressed in terms of range class (Table 1), defined as:
  - Range 0: unanimity among all experts giving a FORTA rating (no deviation);
  - Range 1: greatest range only from A to B, B to C or C to D (neighbouring classes), half weight;



**Table 1** Frequency of substances in defined range groups according to degree of consensus

Range	Frequency ( <i>n</i> = 190)	Percent
0	54	28.42
1	86	45.26
2	43	22.63
3	7	3.68

- Range 2: greatest distance from A to C or B to D, two-thirds weight;
- Range 3: greatest distance from A to D, full weight.

In order to confirm the original/determine new, rater-based labels, we converted the experts' FORTA ratings into numerical values representing the median: A → 1, B → 2, C → 3 and D → 4, respectively. The mean and mode were calculated for each item, reconverted to FORTA labels and compared with the original author-based labels. The range for each label was defined as:

If  $1 \leq m < 1.5 \rightarrow$  FORTA Class A  
 If  $1.5 \leq m < 2.5 \rightarrow$  FORTA Class B  
 If  $2.5 \leq m < 3.5 \rightarrow$  FORTA Class C  
 If  $m \geq 3.5 \rightarrow$  FORTA Class AD

where  $m$  = arithmetic mean based on the raters' grades 1–4.

The scale has not been defined in order to allow for complex statistical calculations. The purpose of this scale was to pool the judgements by taking into account each rater judgement. This was necessary to enable a comparison to be made between the raters' opinions and the FORTA classifications. The assignment of A → 1, B → 2, C → 3 and D → 4 seems to be plausible when assuming that any difference (1–2, 2–3, 3–4) is equally important [31].

5. Round 2: Substances falling short of the preset corrected consensus cutoff of 0.800 were re-sent to the experts. New substances suggested by  $\geq 2$  raters and all new indications were sent to the experts for evaluation in the form of a questionnaire.
6. Analysis based on Round 2 input was performed as follows: confirmation of FORTA labels/determination of new rater-based FORTA labels for re-evaluated items derived from the arithmetic mean (as described above); review of all comments. A simple 'agree-disagree' approach or 5-point scale was not chosen; quantification of disagreement seemed necessary as there are four categories of answers (FORTA classes), and a full match is unlikely to occur; thus, this approach may possibly have led to an equally large second round. In this way, the actual FORTA classes could be preserved as such and either confirmed or challenged.
7. Compilation of an annotated FORTA List based on experts' input over two rounds.

The substances and indications suggested by the experts were selected as follows:

1. Acceptance of all substances/indications receiving an affirmative response by  $>50\%$  of experts during

Round 2 and receiving a FORTA rating (excluding abstentions) by  $\geq 8$  raters.

2. Calculation of a kappa index reflecting label dispersion: here, kappa is defined as the (proportion of 'matching' labels – 0.25)/0.75. This gives due consideration to the fact that a figure of 25 % can theoretically be attained by chance alone, with the choice of four distinct labels.
3. Conversion to median and calculation of mean and mode, as in the first procedure. The arithmetic mean provided the basis for conversion to FORTA labels.
4. Compilation of all substances in a separate, annotated list.

### 3 Results

Twenty experts, 17 geriatric internists and three geriatric psychiatrists representing Germany (13) and Austria (7) agreed to participate in the survey. The return rate for both rounds was 100 %.

The overall consensus for all items and experts after Round 1 was found to be 92 %, corrected (mean 0.922, median 0.950, range 0.500–1.000). Overall, 54/190 (28.4 %) of the evaluated items elicited unanimous agreement among the raters (Table 1) and 24/190 (12.6 %) items fell short of the cutoff of 0.800 and were re-evaluated in a second round (Table 2). Of these 24 items, 19 (79.2 %) represented substances commonly used for the prevention or treatment of dementia and dementia syndromes; 3/24 substances (12.5 %) represented drugs used for treating cardiovascular diseases; 1/24 (4.2 %) was a drug prescribed for depression and 1/24 (4.2 %) for osteoporosis.

Backed by experts' largely convergent comments on the individual substances, two consistent trends could thus be detected. These indicated a shift in the FORTA labels for drugs used to prevent or treat dementia and dementia syndromes:

1. Agreement with the original C label for substances administered for dementia was observed in most cases for the participating geriatric psychiatrists, as opposed to geriatric internists, who tended to favour the D label in these cases. In future FORTA developments, this area should be rated by a larger group of geriatric psychiatrists to emphasize their particular experience in this area.
2. The original D label was challenged specifically for neuroleptic drugs by both geriatric psychiatrists and geriatric internists; many experts tended towards C and expressed the wish for further differentiation/qualifying statements pertaining to the therapy of behavioural and psychological symptoms of dementia (BPSD).

**Table 2** Analysis of the 24 re-evaluated substances

Re-evaluated substance/group (original FORTA rating)	FORTA indication area	No. of raters ( <i>n</i> = 20)		Consensus coefficient Round 1 (cutoff 0.800)	Expert rating on a numerical scale <sup>a</sup> Round 1 Round 2 mean; mode	Proposed FORTA rating, based on mean value from Round 2
		Round 1	Round 2			
<b>Drugs to prevent or treat dementia or dementia-related syndromes</b>						
Nimodipine (C)	Dementia	20		0.750	3.5; 3	D
		19			3.7; 4	
Antioxidants (C)	Dementia	19		0.710	3.6; 4	D
		20			3.9; 4	
Phytotherapeutic agents, e.g. Ginseng (C)	Dementia	20		0.725	3.6; 4	D
		20			3.8; 4	
Hormone preparations, e.g. DHEA (C)	Dementia	20		0.700	3.6; 4	D
		20			3.9; 4	
Gingko biloba (C)	Dementia	20		0.775	3.5; 3	D
		20			3.6; 4	
Ergoline derivatives (C)	Dementia	19		0.763	3.5; 3	D
		20			3.8; 4	
Piracetam (C)	Dementia	20		0.800	3.4; 3	D
		20			3.6; 4	
Pyritinol (C)	Dementia	18		0.778	3.4; 3	D
		19			3.7; 4	
Selegiline (C)	Dementia	19		0.763	3.5; 3	D
		20			3.7; 4	
Haloperidol (D)	BPSD paranoia, hallucinations	19		0.632	3.3; 4	C
		20			3.0; 3	
Risperidone (D)	BPSD paranoia, hallucinations	20		0.500	3.0; 2	C
		20			2.7; 2	
Quetiapine (D)	BPSD paranoia, hallucinations	20		0.575	3.2; 4	C
		20			2.9; 3	
Aripiprazole (D) [2–15 mg/day]	BPSD paranoia, hallucinations	19		0.789	3.6; 4	C
		17			3.4; 4	
Clozapine (D) [10–50 mg/day]	BPSD paranoia, hallucinations	20		0.800	3.6; 4	D
		19			3.7; 4	
Risperidone (D)	BPSD restlessness	20		0.625	3.3; 4	C
		20			2.7; 2	
Melperone (D)	BPSD restlessness	20		0.675	3.4; 4	C
		20			3.4; 4	
Quetiapine (D) [25–200 mg/day]	BPSD restlessness	19		0.763	3.5; 4	C
		19			3.3; 3	
Pipamperone (D) [20–150 mg/day]	BPSD restlessness	19		0.790	3.6; 4	D
		17			3.6; 4	
Mirtazapine (C) [15–30 mg]	BPSD sleep disorders	20		0.775	3.0; 3	C
		20			3.0; 3	
Mirtazapine (D)	Insomnia	20		0.700	3.4; 4	C
		20			3.45; 4	
<b>Other drugs</b>						
Digitoxin (D)	Atrial fibrillation	20		0.525	3.1; 4	C
		19			2.5; 2	
Digoxin (B)	Atrial fibrillation	20		0.800	2.4; 2	B
		19			2.4; 2	

**Table 2** continued

Re-evaluated substance/group (original FORTA rating)	FORTA indication area	No. of raters ( <i>n</i> = 20)		Consensus coefficient Round 1 (cutoff 0.800)	Expert rating on a numerical scale <sup>a</sup> Round 1 Round 2 mean; mode	Proposed FORTA rating, based on mean value from Round 2
		Round 1	Round 2			
Dronedarone (B)	Atrial fibrillation	18		0.556	2.9; 3	C
		18				
Strontium ranelate (B)	Osteoporosis	17		0.794	2.1; 2	B
		18				

FORTA Fit FOR The Aged, DHEA dehydroepiandrosterone, BPSD behavioural and psychological symptoms of dementia

*A* → 1; *B* → 2

*C* → 3; *D* → 4

<sup>a</sup> If  $1 \leq m < 1.5 \rightarrow A$

If  $1.5 \leq m < 2.5 \rightarrow B$

If  $2.5 \leq m < 3.5 \rightarrow C$

If  $m \geq 3.5 \rightarrow D$

Due to abstentions, the number of raters varied for each item (maximum 20, range 5–20). According to the raters' comments, the most common reason for abstaining was insufficient experience with a particular substance or indication group and, in individual cases, potential conflict of interest.

The indication area of oncological/haematological illnesses (27 items) received the most abstentions and thus the lowest number of raters (mean 7.41, median 7.22, range 5–12), yet the calculated corrected consensus values were consistently high (mean 0.964, range 0.857–1.000) for all items tested. Most experts gave the reason for abstention as insufficient experience or lack of familiarity with the current state of evidence for oncological treatments. One rater documented a consultation with other experts in the field of oncology. Re-evaluation was foregone, but this indication group will be kept under close scrutiny during further clinical development.

According to the calculations based on a numerical scale for purposes of comparing the rater-based labels to the original author-based labels, it was found that, after Round 1, 12 of the 24 re-evaluated items (6.3 % of the original 190 substances) had received a FORTA label diverging from the original label. These 12 items correlated directly with the substances receiving the lowest corrected consensus coefficients. After Round 2, this number had increased to 19 of the 24 retested substances (5/24 labels confirmed), eliciting final confirmation of 90 % (171/190) of the original labels. This increase in label deviation appears to indicate that other factors may have played a prominent role in the decision-making process for labelling during Round 2.

A total of 35 new substances were accepted for potential incorporation into the FORTA List. Nineteen substances

were included in four new indication areas: epilepsy (12), anaemia (4), gastrointestinal illnesses/concomitant application of non-steroidal anti-inflammatory drugs [NSAIDs] (2) and bipolar disorder (1). Thus, the original opinion-based proposal of indication areas by the authors of reference [25] was largely confirmed by the rater panel. Sixteen substances associated with pre-existing FORTA indications were included: drugs for the therapy of depression (6), chronic pain (3), atrial fibrillation (2), arterial hypertension (2), coronary heart disease (1), osteoporosis (1) and insomnia (1).

These results are summarized in the FORTA List, available online as ESM 1 (full statistical details including all results from the first round are available upon request).

#### 4 Discussion

The Delphi method often presents a challenge to carry out in practice, not least due to the lack of evidence in the literature as to optimal standard operation procedures and forms of interpretation. Nevertheless, it has become an acceptable mode, and sometimes the only feasible option, of obtaining experts' opinions on particularly complex topics [27, 32, 33].

Medication lists and classification systems developed during the past few decades represent a variation on an established theme. In the Comprehensive Drug Abuse and Control Act of 1970, for example, harmful or habit-forming drugs were ranked according to their 'dangerousness'; selected drugs were assigned to 'specific ... categories with appropriate restrictions' [34]. Our proposed FORTA system involves the evidence-based classification of medications according to age-appropriateness. Through the expert

validation procedure, the FORTA List, a drug-appropriateness rating system, has been endorsed and improved by the input of 20 experts, thereby enhancing its value for implementation in a clinical setting, while areas requiring further attention and development clearly came into view. The panel was chosen from clinical specialties only as clinical experiences in the elderly appeared to be most valuable in this patient population, which affected decisions made as to the choice of inclusion of other specialties (e.g. pharmacoepidemiology, pharmacists).

The still relatively tenuous or inconsistent state of evidence associated with medications for dementia is also reflected in examples from the literature [35–37]. This area would appear to require further observation and development during clinical studies. Future clinical projects will also specifically have to target the problem of how best to classify dementia in the FORTA List. Many experts discussed possible benefits of classification according to etiology (i.e. Alzheimer's vs. vascular origin). The FORTA List divides dementia into subclasses according to additional syndromes (BPSD); drug therapy is either of preventive nature or symptom-oriented and has been simplified here to the greatest extent possible. More specific differentiation during further clinical studies may improve the overall quality and practicability of the system.

The first positive indications of FORTA's potential usefulness in everyday clinical routine are apparent in results obtained from a pilot study applying the FORTA principle [38] to the drug therapy of 46 patients in a geriatric clinic in Essen, Germany. It could be demonstrated that the number of Class A and B medications significantly increased, and the number of Class C and D medications were reduced. Preliminary data obtained from a prospective, single blinded, randomized trial involving 97 patients, also conducted in Essen, further revealed that use of FORTA may be associated with a reduction in in-hospital falls [39].

Further-reaching applications of the FORTA system may include refining the process of defining and assigning FORTA labels (classes A–D) to newly selected and already-established drugs assessed by Health Technology Assessment institutions [for example, the National Institute for Health and Care Excellence (NICE) in England, or the Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG, Institute for Quality and Economic Effectiveness in Health Care in Germany)]. Although as yet lacking in most major situations, future input provided by controlled and, ultimately, real-life studies represent equally essential components of the development procedure for this drug labelling system, since adherence factors, availability and application issues also play important roles in determining the ultimate effectiveness and safety of these

substances [40]. Another future task is the further differentiation and separation of distinct compounds that now reside in a 'mixed pot', e.g. the group of 'frequency-lowering  $\beta$ -blockers' (FORTA list, p. 11), which formally still contains sotalol but which is now considered mainly as a class III antiarrhythmic (D drug), or propranolol, which should not be used (with exceptions) for pharmacokinetic reasons. For treatment of heart failure with  $\beta$ -blockers, positively labelled compounds are listed (p. 10). Similarly, not all dihydropyridines are well studied in the elderly, and amlodipine is given as a lead example in the FORTA list (p. 9), reflecting the results of the Avoiding Cardiovascular Events in Combination Therapy in Patients Living with Systolic Hypertension (ACCOMPLISH) and Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT) trials [41, 42]. Nor are diuretics further differentiated, although long-term data are mainly available for thiazides, e.g. in ACCOMPLISH. The choice between loop and thiazide diuretics is guided by renal function and/or severity of heart failure, but does not lead to different assessments as yet, mainly due to the similarity in side effects (e.g. electrolyte disorders).

If compared with START/STOPP criteria, drugs recommended by START seem to belong predominantly to FORTA classes A or B. Examples include statins or acetylsalicylic acid in the treatment of coronary heart disease, ACE inhibitors for heart failure, levodopa for Parkinson's disease, or corticosteroids and inhaled  $\beta_2$ -agonists in chronic obstructive pulmonary disease. Substances mentioned in the STOPP list [20] appear to correspond with categories C and D as assigned by the FORTA system (see also revised FORTA List). Here, examples include benzodiazepines—unanimously voted across all systems to be potentially inappropriate or negatively rated, either in drug–disease connection or in and of themselves [17, 18, 20, 25, 43]—neuroleptic drugs, first-generation antihistamines or theophylline as monotherapy in chronic obstructive pulmonary disease.

Not unexpectedly for consensus processes, a few discrepancies with the updated Beers list are also present. Although difficult to compare directly, as individual substances are taken into consideration in relation to specific illnesses or conditions (drug–disease aspect) [17, 43], one notable example here includes the classification or use/non-use of digoxin, which is listed to be avoided in higher dosages. Concordance between Beers and FORTA is however high; specific examples include the mention of NSAIDs to be avoided in chronic use, doxazosin to avoid as an antihypertensive, benzodiazepines and zolpidem to avoid in most instances, and carbamazepine to be used with caution, corresponding to the FORTA classes C and D for these selected substances.

Compared to the PRISCUS list, most PIMs have been labelled C or D, with few exceptions, most notably digoxin,

as already mentioned above (FORTA B, PIM in PRISCUS, Beers and STOPP) [17, 18, 20, 43]. This compound and its congener, digitoxin, are on the list with major discrepancies in ratings (Table 2); digoxin was rated favourably in the treatment of atrial fibrillation as renal dosing is amenable and intoxication effects are much shorter than for digitoxin, still prescribed in Germany and rated FORTA C.

A major advance made by development of the FORTA system has led to the quantitative assessment allowing for cross-therapeutic prioritization and reflection of multiple diseases leading to reduced medication schemes, whereas START/STOPP criteria could still lead to additive polypharmacy if multiple conditions are met. The user is introduced to a standard, reproducible system, the repeated employment of which may encourage an overall learning effect ('geriatric pharmacology in a nutshell').

The 'internationalization' of the FORTA List may be viewed as one of the next important steps in the development of FORTA. In this context it is important to acknowledge that most PIM lists and clinical tools do remain country-specific, both in Europe (most European countries such as Germany, France, Norway and Austria have their own negative listings) and in the US (Beers List). This reflects the diversity of national drug use and regulatory status. It may however be noted that the original authors did not encounter any major problems converting to the US system, as documented in the first English-language edition of the original FORTA source containing the author-based, US version of the FORTA List [44]. Although not developed for European countries, the Beers and McLeod lists from the US and Canada, respectively, have been successfully used to detect and compare PIMs in eight European countries [45]. Thus, drug listing approaches seem to be principally applicable even to geographically removed industrialized societies. Yet, the well-known potential obstacles of divergence and differences in drug availability, as well as country-specific prescribing trends, demography and disease epidemiology must not be ignored. Thus, ensuing 'gaps' or inconsistencies, while not actively presenting a hindrance in our estimation, still represent an area requiring intensified cooperative efforts, ideally on an international level.

## 5 Limitations of the Survey

Important limitations of the Delphi process which arose and should be mentioned here include the following issues:

1. The choice of raters did not include a wide array of experts, e.g. general practitioners for ambulatory care, pharmacists or higher numbers of geriatric psychiatrists.
2. The FORTA List may have limited applicability for international use and still awaits adaptations in an internationalization process.
3. There is a relative lack of evidence-driven ratings compared with consensus-driven ratings. Future modifications should emphasize evidence-driven ratings, particularly in emerging areas of age-related therapeutic knowledge or innovations, e.g. novel oral anticoagulants.
4. Due to the emphasis on implicit criteria (the individual patient has to be considered for the application of the FORTA List) the utility of the FORTA tool may be limited regarding its use in pharmacoepidemiological research.
5. FORTA does not specifically address drug–drug interactions or contraindications which still need to be checked individually, as well as drug doses and medication scheduling; it does not aim at detecting prescribing cascades, although an increased quality of prescriptions will certainly help to reduce them.

## 6 Conclusion

When applied according to specific, well-defined criteria within the context of individualized patient care and management, the FORTA List should help physicians to optimize drug treatment in their older patients. The expert consensus validation process for the FORTA List was essential in its development, and it is our hope that this will ultimately facilitate its use in clinical practice.

**Acknowledgments** This study was supported by a grant from the Deutsche Forschungsgemeinschaft (WE 1184/15-1), which sponsored and approved the planning, execution and analysis as proposed in the grant application. We are grateful to Christine Schummer for her secretarial support.

Martin Wehling was responsible for the creation and development of the original FORTA concept, as well as the overall supervision and approval of all contents. Alexandra Kuhn-Thiel coordinated and oversaw the Delphi consensus procedure and contributed to the design and organization of the revised FORTA List. All aspects of this process were performed under close team monitoring. Christel Weiß performed all statistical analysis for the Delphi procedure. All participating experts reviewed and approved the manuscript and contributed actively to the content and presentation of the FORTA List.

Martin Wehling was employed by AstraZeneca R&D, Mölndal, as Director of discovery medicine (= translational medicine) from 2003 to 2006, while on sabbatical leave from his professorship at the University of Heidelberg. Since returning to this position in January 2007 he has received lecturing and consulting fees from Sanofi-Aventis, Novartis, Takeda, Roche, Pfizer, Bristol-Myers, Daichii-Sankyo, Lilly, Novo-Nordisk, Shire and LEO Pharma. Christel Weiß and Alexandra Kuhn-Thiel have no conflicts of interest to declare.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## Appendix

### Original FORTA List Authors<sup>[25]</sup> and Their Affiliations

*Martin Wehling, MD:* Institute for Experimental and Clinical Pharmacology and Toxicology, Department of Clinical Pharmacology, Center for Geriatric Pharmacology, Medical Faculty Mannheim, University of Heidelberg, Maybachstr. 14, 68169 Mannheim, Germany

*Heinrich Burkhardt, MD:* University Medical Centre Mannheim, 4th Department of Medicine, Geriatric Medicine, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany

*Lutz Frölich, MD:* Central Institute of Mental Health, J5, 68159 Mannheim, Germany

*Stefan Schwarz, MD:* Central Institute of Mental Health, J5, 68159 Mannheim, Germany

*Ulrich Wedding, MD:* University Hospital, Clinic for Internal Medicine II, Erlanger Allee 101, 07747 Jena, Germany

### Expert Panel Members and Their Affiliations

*Jürgen Bauer, MD:* Geriatrics Centre Oldenburg, University of Oldenburg, Rahel-Straus-Straße 10, 26133 Oldenburg, Germany

*Heiner K. Berthold, MD:* Charité University Medicine Berlin, Evangelical Geriatrics Center Berlin (EGZB), Reinickendorfer Str. 61, 13347 Berlin, Germany

*Peter Dovyjak, MD:* Gmunden Hospital, Department of Acute Geriatric Medicine, Miller-von-Aichholz-Straße 49, A-4810 Gmunden, Austria

*Helmut Frohnhofen, MD:* Essen-Mitte Hospital, Knappschafts Hospital, Teaching Hospital, University of Duisburg-Essen, Am Deimelsberg 34a, 45276 Essen, Germany, and Faculty of Health, University of Witten-Herdecke, Witten, Germany

*Thomas Frühwald, MD:* Hietzing Hospital and Neurological Center Rosenhügel, Wolkersbergenstraße 1, 1130 Vienna, Austria

*Christoph Gisinger, MD:* Haus der Barmherzigkeit, Danube University Krems, Seeböckgasse 30a, 1160 Vienna, Austria

*Manfred Gogol, MD:* Lindenbrunn Hospital, Geriatric Department, Lindenbrunn 1, 31863 Copenbruegge, Germany

*Markus Gosch, MD:* Regional Hospital Hochzirl, Anna-Dengel House, 6170 Zirl, Austria

*Hans Gutzmann, MD:* Hedwigshöhe Hospital, Clinic for Psychiatry, Psychotherapy and Psychosomatic Medicine, Höhensteig 1, 12526 Berlin, Germany

*Isabella Heuser, MD:* Charité University Hospital Berlin, Department of Psychiatry and Psychotherapy, University Medicine Berlin, Campus Benjamin Franklin, Eschenallee 3, 14050 Berlin, Germany

*Werner Hofmann, MD:* Friedrich Ebert Hospital, Clinic for Geriatric Medicine, Friesenstrasse 11, 24534 Neumünster, Germany

*Michael Hüll, MD:* Center for Geriatric Medicine and Gerontology Freiburg, University Clinic Freiburg, Lehener Straße 88, 79106 Freiburg, Germany

*Bernhard Iglseder, MD:* Department of Geriatric Medicine, Christian-Doppler-Klinik, Paracelsus Medical University, Ignaz-Harrer-Str. 79, 5020 Salzburg, Austria

*Anja Kwetkat, MD:* Jena University Hospital, Department of Geriatric Medicine, Bachstraße 18, 07743 Jena, Germany

*Michael Meisel, MD:* Deaconess Hospital Dessau, Clinic for Internal and Geriatric Medicine, Gropiusallee 3, 06846 Dessau, Germany

*Wolfgang Mühlberg, MD:* Clinic for Internal Medicine 4 – Geriatric Medicine, Frankfurt Höchst Hospital, Gote-nstraße 6-8, 65929 Frankfurt am Main, Germany

*Wolfgang von Renteln-Kruse, MD:* Albertinen Hospital/ Albertinen House, non-profit company (GmbH), Center for Geriatric Medicine and Gerontology, Scientific Institution at the University of Hamburg, Sellhopsweg 18-22, 22459 Hamburg, Germany

*Regina Roller, MD:* Medical University of Graz, Department of Internal Medicine, Auenbruggerplatz 15, 8036 Graz, Austria

*Ralf-Joachim Schulz, MD:* Geriatric Clinic at the St.-Marien Hospital, Kunibertkloster 11-13, 50668 Köln, Germany

*Ulrike Sommeregger, MD:* Hietzing Hospital and Neurological Center Rosenhügel, Wolkersbergenstraße 1, 1130 Vienna, Austria.

## References

1. US Department of Commerce, US Census Bureau. The older population; 2011 (online). <http://www.census.gov/population/age/>. Accessed 2 Jul 2013.
2. Böhm K, Tesch-Römer C, Ziese T (eds) Gesundheitsberichterstattung des Bundes 2009: Gesundheit und Krankheit im Alter. Berlin: Robert Koch Institute; 2009 [in German]. [http://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/alter\\_gesundheit.pdf?\\_\\_blob=publicationFile](http://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/alter_gesundheit.pdf?__blob=publicationFile). Accessed 2 Jul 2013.
3. Kaufman DW, Kelly JP, Rosenberg L, et al. Recent patterns of medication use in the ambulatory adult population of the United States: the Slone Survey. *JAMA*. 2002;287:337–44.
4. Van den Akker M, Buntinx F, Metsmakers JF, et al. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J Clin Epidemiol*. 1998;51:367–75.
5. Mangoni AA, Jackson SHD. Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. *Br J Clin Pharmacol*. 2003;57:6–14.

6. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet*. 2000;356:1255–9.
7. Rochon PA, Gurwitz JH. Optimising drug treatment for elderly people: the prescribing cascade. *Br Med J*. 1997;315:1096–9.
8. Chyka PA. How many deaths occur annually from adverse drug reactions in the United States? *Am J Med*. 2000;109:122–30.
9. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*. 1998;279:1200–5.
10. Borchelt M. Potential side-effects and interactions in multiple medication in elderly patients: methodology and results of the Berlin Study of Aging. *Z Gerontol Geriatr*. 1995;28:420–8.
11. Michel J-P, Pils K, Sieber C. Commentary on “The future history of geriatrics: geriatrics at the crossroads”. *J Gerontol Med Sci*. 2002;57A:M812–3.
12. Evans JG. Geriatric medicine: a brief history. *Br Med J*. 1997;315:1075–7.
13. Wehling M. Guideline-driven polypharmacy in elderly, multimorbid patients is basically flawed: there are almost no guidelines for these patients. *J Am Geriatr Soc*. 2011;59:376–7.
14. Gillick M. Choosing appropriate medical care for the elderly. *J Am Med Dir Assoc*. 2001;2:305–9.
15. Barenholtz Levy H, Marcus EL, Christen C. Beyond the Beers Criteria: a comparative overview of explicit criteria. *Ann Pharmacother*. 2010;44:1968–75.
16. Beers MH. Explicit criteria for determining potentially inappropriate medication use by the elderly: an update. *Arch Intern Med*. 1997;157:1531–6.
17. Fick DM, Semla TP. American Geriatrics Society Beers Criteria: new year, new criteria, new perspective. *J Am Geriatr Soc*. 2012;60:614–5.
18. Holt S, Schmiedl S, Thürmann PA. Potentially inappropriate medications in the elderly: the Priscus List. *Dtsch Arztebl Int*. 2010;107:543–51.
19. Page RL, Ruscin JM. The risk of adverse drug events and hospital-related morbidity and mortality among older adults with potentially inappropriate medication use. *Am J Geriatr Pharmacother*. 2006;4:297–305.
20. 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.
21. Barry PJ, Gallagher P, Ryan C, O'Mahony D. START (screening tool to alert doctors to the right treatment): an evidence-based screening tool to detect prescribing omissions in elderly patients. *Age Ageing*. 2007;36:632–8.
22. Lam MPS, Cheung BMY. The use of STOPP/START criteria as a screening tool for assessing the appropriateness of medications in the elderly population. *Expert Rev Clin Pharmacol*. 2012;5:187–97.
23. Wehling M. Drug therapy in the elderly: too much or too little, what to do? A new assessment system: fit for the aged (FORTA). *Dtsch Med Wochenschr*. 2008;133:2289–91 [in German].
24. Wehling M. Multimorbidity and polypharmacy: how to reduce the harmful drug load and yet add needed drugs in the elderly? Proposal of a new drug classification: fit for the aged. *J Am Geriatr Soc*. 2009;57:560–1.
25. Wehling M, Burkhardt H. *Arzneitherapie für Ältere*. 2nd ed. Heidelberg: Springer; 2011 [in German].
26. Hsu C, Sandford BA. The Delphi technique: making sense of consensus. *Pract Assess Res Eval* 2007;12:1–7. <http://pareonline.net/getvn.asp?v=12&n=10>. Accessed 2 Jul 2013.
27. Häder M. *Delphi-Befragungen: Ein Arbeitsbuch*. 2nd ed. Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH; 2009 [in German].
28. Cicchetti DV, Allison T. A new procedure for assessing reliability of EEG sleep recordings. *Am J EEG Technol*. 1971;11:101–9.
29. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70:213–20.
30. Stevens SS. On the theory of scales of measurement. *Science*. 1946;103(2684):677–80. <http://www.jstor.org/stable/1671815>. Accessed 12 Oct 2013.
31. Altman DG. *Practical statistics for medical research*. London: Chapman & Hall; 1990. p. 15.
32. Fitch K, Bernstein S, Aguilar MD, et al. *The RAND/UCLA appropriateness method user's manual*. Santa Monica: RAND; 2001.
33. Stitt-Gohdes WL, Crews TB. The Delphi technique: a research strategy for career and technical education. *JCTE*. 2004;20:55–67.
34. Musto DF. *The American disease: origins of narcotic control, expanded edition*. New York: Oxford University Press, Inc; 1987.
35. Alexopoulos GS, Streim J, Carpenter D, Docherty JP, Expert Consensus Panel for Using Antipsychotic Drugs in Older Patients. Using antipsychotic agents in older patients. *J Clin Psychiatry*. 2004;65 Suppl 2:5–99.
36. Stevenson DG, Decker SL, Dwyer LL, et al. Antipsychotic and benzodiazepine use among nursing home residents: findings from the 2004 National Nursing Home Survey. *Am J Geriatr Psychiatry*. 2010;18:1078–92.
37. Birks J, Grimley Evans J. Ginkgo biloba for cognitive impairment and dementia. *Cochrane Database Syst Rev*. 2009;(1):CD003120.
38. Frohnhofen H, Michalek C, Wehling M. Assessment of drug treatment in geriatrics with the new FORTA criteria: preliminary clinical experience. *Dtsch Med Wochenschr*. 2011;136:1417–21 [in German].
39. Frohnhofen H, Wehling M, Michalek C. Erste klinische Erfahrungen mit der neuen FORTA-Klassifikation (Abstract). *Z Gerontol Geriatr*. 2012;45:21 [in German].
40. Field TS, Mazor KM, Briesacher B. Adverse drug events resulting from patient errors in older adults. *J Am Geriatr Soc*. 2007;55:271–6.
41. Jamerson K, Bakris GL, Dahlöf B, et al. Exceptional early blood pressure control rates: the ACCOMPLISH Trial. *Blood Pressure*. 2007;16:80–6.
42. Moser M. The Ascot Trial. *J Clin Hypertens*. 2005;7:748–50.
43. American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc*. 2012;2012(60):616–31.
44. Wehling M. *Drug therapy for the elderly*. Vienna: Springer; 2013.
45. Fialová D, Topinková E, Gambassi G, AdHOC Project Research Group, et al. Potentially inappropriate medication use among elderly home care patients in Europe. *JAMA*. 2005;293:1348–58.