

# Consensus Recommendations for Systematic Evaluation of Drug–Drug Interaction Evidence for Clinical Decision Support

Richard T. Scheife · Lisa E. Hines · Richard D. Boyce · Sophie P. Chung · Jeremiah D. Momper · Christine D. Sommer · Darrell R. Abernethy · John R. Horn · Stephen J. Sklar · Samantha K. Wong · Gretchen Jones · Mary L. Brown · Amy J. Grizzle · Susan Comes · Tricia Lee Wilkins · Clarissa Borst · Michael A. Wittie · Daniel C. Malone

Published online: 3 January 2015  
© Springer International Publishing Switzerland 2015

## Abstract

**Background** Healthcare organizations, compendia, and drug knowledgebase vendors use varying methods to evaluate and synthesize evidence on drug–drug interactions (DDIs). This situation has a negative effect on electronic prescribing and medication information systems that warn clinicians of potentially harmful medication combinations. **Objective** The aim of this study was to provide recommendations for systematic evaluation of evidence for DDIs from the scientific literature, drug product labeling, and regulatory documents.

**Methods** A conference series was conducted to develop a structured process to improve the quality of DDI alerting

systems. Three expert workgroups were assembled to address the goals of the conference. The Evidence Workgroup consisted of 18 individuals with expertise in pharmacology, drug information, biomedical informatics, and clinical decision support. Workgroup members met via webinar 12 times from January 2013 to February 2014. Two in-person meetings were conducted in May and September 2013 to reach consensus on recommendations. **Results** We developed expert consensus answers to the following three key questions. (i) What is the best approach to evaluate DDI evidence? (ii) What evidence is required for a DDI to be applicable to an entire class of drugs? (iii) How should a structured evaluation process be vetted and validated?

**Conclusion** Evidence-based decision support for DDIs requires consistent application of transparent and systematic methods to evaluate the evidence. Drug compendia and

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s40264-014-0262-8) contains supplementary material, which is available to authorized users.

---

R. T. Scheife  
Tufts University School of Medicine, Boston, MA, USA

L. E. Hines · M. L. Brown · A. J. Grizzle · D. C. Malone (✉)  
University of Arizona College of Pharmacy, 1295 N Martin Ave,  
PO Box 210202, Tucson, AZ 85721-0202, USA  
e-mail: malone@pharmacy.arizona.edu

R. D. Boyce  
Department of Biomedical Informatics, University of Pittsburgh,  
Pittsburgh, PA, USA

S. P. Chung · G. Jones · S. Comes  
Eprocrates, athenahealth, Inc., San Francisco, CA, USA

J. D. Momper  
University of California, San Diego, Skaggs School of Pharmacy  
and Pharmaceutical Sciences, La Jolla, CA, USA

C. D. Sommer  
FDB (First Databank, Inc.), South San Francisco, CA, USA

D. R. Abernethy  
Office of Clinical Pharmacology, U.S. Food  
and Drug Administration, Silver Springs, MD, USA

J. R. Horn  
University of Washington School of Pharmacy,  
Seattle, WA, USA

S. J. Sklar  
Wolters Kluwer Health, Indianapolis, IN, USA

S. K. Wong  
Cerner Multum, Denver, CO, USA

T. L. Wilkins · M. A. Wittie  
U.S. Office of the National Coordinator for Health  
Information Technology, Washington, DC, USA

C. Borst  
Elsevier Clinical Solutions, Tampa, FL, USA

clinical decision support systems in which these recommendations are implemented should be able to provide higher-quality information about DDIs.

### Key Points

Evidence-based clinical decision support for drug–drug interactions (DDIs) requires consistent application of transparent and systematic methods to evaluate the evidence.

An expert workgroup developed recommendations by consensus for systematic evaluation of evidence for DDIs from the scientific literature, drug product labeling, and regulatory documents.

Workgroup members expect that editors of drug compendia and clinical decision support systems will be able to provide higher-quality information about DDIs.

## 1 Background

Exposure to potential drug–drug interactions (DDIs) is a significant source of preventable drug-related harm that requires proper management to avoid medical errors [1]. Studies indicate DDIs harm 1.9–5 million inpatients per year and cause 2,600–220,000 emergency department visits in the US per year [2–4].

The importance of DDIs as a risk factor for patient harm led the Centers for Medicare and Medicaid Services (CMS) to include DDI clinical decision support (CDS) alerts in the agency’s guidelines for achieving meaningful use of electronic health records [5]. However, evidence indicates that DDI decision support systems have not successfully reduced exposure to DDIs [6–8]. In the US, most alerting systems rely on clinical content created, maintained, and sold by knowledgebase vendors [9]. Each organization implements their own approach to classifying DDIs with limited agreement between systems [10–12]. Additionally, CDS systems often alert for DDIs that have limited clinical relevance, which may increase alert fatigue [13] and lead to inappropriate responses [14–16].

In spite of a desire among providers of DDI decision support tools to produce clinically relevant content, improving the state-of-the-art poses several challenges. High-quality evidence to support the existence of many DDIs is lacking, there are few controlled clinical studies

conducted in relevant populations [17–19], and individual case reports are underreported and often lack information [20]. Compendia and knowledgebase editors use differing approaches to identify and evaluate evidence [10–12]. There are no guidelines or standards for determining clinical relevance of interactions via consistent systematic evaluation or classification [9, 21]. Without such guidance, DDIs may also be inappropriately extrapolated to other drugs within the same therapeutic or pharmacologic class [22]. In an effort to reduce legal liability, system vendors might have an incentive to include almost all possible DDIs, including those that confer extremely low risk to exposed patients [9, 23].

A conference series was conducted to develop specific recommendations to improve the quality of CDS alerts for DDIs. The goals of the conference series were addressed by three workgroups. This paper describes recommendations by the Evidence Workgroup that was charged with the objective to provide recommendations for systematic evaluation of evidence from the scientific literature, drug product labeling, and regulatory documents with respect to DDIs for CDS. These activities were supported in part by a conference grant from the Agency for Healthcare Research and Quality (AHRQ) and donations from health information technology (IT) vendors. Use of funds was at the sole discretion of the University of Arizona and according to Department of Health and Human Services requirements.

## 2 Methods

Nineteen individuals (listed as co-authors) with expertise in DDIs, clinical pharmacology, drug information, evidence evaluation, biomedical informatics, and health IT were invited to participate as workgroup members. No invited experts declined to participate; one individual was unable to contribute due to health reasons and is not listed as a co-author. Members represented diverse backgrounds such as academia; journal, compendia, and knowledgebase editors; healthcare organizations; US Food and Drug Administration (FDA); and the US Office of the National Coordinator for Health IT (ONC). Twelve 1-h webinar meetings were conducted from January 2013 to February 2014, with two in-person meetings held in Washington DC (May 2013) and Phoenix, Arizona (September 2013). A member with recognized leadership skills and experience facilitated the meetings. The group followed a structured consensus-development process that included clarifying issues to be decided; open discussion and debate among all members; iterative aggregation and refinement of ideas leading to proposals that incorporated the best elements of members’ ideas while addressing all key concerns; and active agreement on the final proposal. To ensure the most pressing

issues were addressed, a nonsystematic search of the literature was conducted for papers describing methods for evaluating DDI evidence. From these articles, the following key questions were developed by the conference organizers and then reviewed and agreed upon by consensus of the members: (i) What is the best approach to evaluate DDI evidence? (ii) What evidence is required for a DDI to be applicable to an entire class of drugs? (iii) How should a structured evaluation process be vetted and validated?

Workgroup members were provided access to articles that were deemed relevant for consideration. They also identified relevant studies and copies were obtained for all workgroup members to review. Each key question was evaluated in light of the available evidence and the collective experience of the workgroup members. Responses to each key question were written and then modified to improve clarity or address issues or concerns. Workgroup recommendations were posted on a project Internet site and feedback was sought from other stakeholders via dissemination to professional societies and organizations. Consensus was achieved through an iterative process of drafting recommendations, collecting verbal or written comments from workgroup members and other content experts, and revising documents until no additional substantive comments were provided. Workgroup recommendations were presented at regional and national forums to solicit feedback from stakeholders such as healthcare providers, compendia and knowledgebase editors, professional and quality organizations, and government agencies. Members of the workgroup were informed of feedback during the regularly scheduled webinar meetings. No substantial changes were made to the recommendations based on comments and questions collected during this vetting process. Changes to the recommendations were editorial in nature to improve clarity.

### 3 Results

A summary of our recommendations when evaluating the DDI evidence is shown in Table 1. Details about these recommendations are described more fully below. The results are presented in the following order: (i) recommendations about terminology; (ii) best approaches for evaluating DDI evidence (Key Question 1); (iii) recommendations for evidence of drug-class interactions (Key Question 2); and (iv) procedures to validate a structured process for DDI evidence evaluation (Key Question 3).

#### 3.1 Terminology

We recommend consistent use of relevant terminology for evaluation of DDI evidence. In the process of answering

the key questions, several terms required clarification. A complete list of definitions agreed upon by the workgroup is provided in Electronic Supplementary Material (ESM) 1, with some key terms described below.

A DDI is defined as a clinically meaningful alteration in the exposure and/or response to a drug (object drug) that has occurred as a result of the co-administration of another drug (precipitant drug) [24, 25]. Response can refer to either precipitating an adverse event or altering the therapeutic effect of the object drug. A *potential DDI* is defined as the co-prescription of two drugs known to interact, and therefore a DDI could occur in the exposed patient [25]. Although the distinction between a DDI and a potential DDI is important, both are referred to as DDI throughout this paper for simplicity. A *clinically relevant DDI* is defined as one associated with either toxicity or loss of efficacy that warrants the attention of healthcare professionals. We recommend use of the term *seriousness*, rather than *severity*, to describe the extent to which a DDI can or does cause harm [26].

We developed a working definition for narrow therapeutic index (NTI) because many clinically relevant DDIs involve drugs with a NTI. Similar terms include narrow therapeutic ratio and narrow therapeutic range (NTR). Existing NTI/NTR definitions (see ESM 2) were considered inadequate for guiding the evaluation of DDIs [27]. Although the FDA is developing a definition of NTI drugs in the context of bioequivalence, this definition would generally be stricter than is needed for managing DDIs in clinical practice. Therefore, we define NTI drugs as those for which even a small change in drug exposure may lead to toxicity or loss of efficacy. Several scenarios describe what may constitute a ‘small’ change in drug levels: <100 % (<2-fold) increase in area under the concentration–time curve (AUC) for the object drug may lead to serious adverse events; or <50 % decrease in AUC for the object drug may result in a loss of efficacy with serious therapeutic consequences (e.g., failure of contraception, or virologic failure due to subtherapeutic drug levels).

#### 3.2 Key Question 1: What is the Best Approach to Evaluate Drug–Drug Interaction (DDI) Evidence?

The first step in evaluating DDIs to guide clinical decision making relates to establishing sufficient evidence that a DDI exists, followed by questions of clinical relevance and how to present DDI recommendations to health professionals. Our recommendations primarily focus on identifying the best approach to evaluate if a DDI exists, with additional consideration for how to establish clinical relevance.

**Table 1** Summary of Evidence Workgroup recommendations for systematic evaluation of DDI evidence

| Recommendation   | Comments  |
|--|---|
| Recommendations to be adopted in the short term  |   |
| Apply consistent terminology   | Terminology defined by the Evidence Workgroup is provided in ESM 1  |
| Apply DIPS for evaluating DDI case reports [20]  | DIPS guides users to conduct an explicit, transparent, complete, and balanced assessment of the attributes important to causality assessment of whether a DDI occurred and exists [29]  |
| Develop a new approach for evaluating a body of evidence   | DRIVE was developed based upon important concepts from existing methods [30, 31] with modification to (i) use simple evidence categories; (ii) include causality assessment with DDI case reports (via DIPS); (iii) apply reasonable extrapolation, including from in vitro studies; (iv) address evidence/statements provided in product labeling; and (v) describe study quality criteria and interpretation in the context of DDIs |
| Evaluate statements/evidence in FDA documents and product labeling by same criteria as published evidence  | DDI recommendations provided in labeling that are not supported by PK data or PD properties of the drugs are insufficient evidence<br>DDI listings/recommendations in CDS systems do not need to align with unsupported statements in product labeling  |
| Classify DDIs by therapeutic/pharmacologic class only when the evidence applies, or can be reasonably extrapolated, to the entire class of drugs | Class effects distinction commonly apply to PD DDIs, but rarely to PK DDIs  |
| Recommendations for future work  |   |
| Evaluate DRIVE Instrument  | Testing should include (i) usability; (ii) inter-rater agreement, (iii) concordance (where expected) with similar systems [30, 31]; and (iv) evaluation of DIPS for case reports [20]   |
| Develop systematic DDI search criteria   | See examples by Boyce et al. [32–34]<br>Search should include unpublished literature, including FDA preapproval reviews and post-marketing analyses [35, 36]<br>Search criteria are needed to identify evidence for (i) the existence of a DDI; (ii) clinical outcomes, including frequency and modifying factors; and (iii) clinical management (e.g., monitoring, dosage adjustments, and therapeutic alternatives)                 |
| Identify or develop a system to evaluate and communicate the strength of DDI evidence with graded recommendations [48]                           | The GRADE approach is endorsed or used by numerous organizations (e.g., AHRQ, Cochrane Collaboration, WHO) [53–56]<br>A modified GRADE approach is used by the University of Liverpool for evidence of DDIs involving medications for HIV and hepatitis C [57, 58]  |

AHRQ Agency for Healthcare Research and Quality, CDS clinical decision support, DDI drug–drug interaction, DIPS Drug Interaction Probability Scale, DRIVE DRug Interaction eVidence Evaluation, ESM electronic supplementary material, FDA US Food and Drug Administration, GRADE Grading of Recommendations, Assessment, Development, and Evaluation, PD pharmacodynamic, PK pharmacokinetic, WHO World Health Organization

When publishing a recommendation about the risk of a DDI, it is essential first to assess the quality of individual studies to prevent drawing erroneous conclusions about the entire body of evidence. Evaluation of medical treatments commonly involves hierarchical rating schemes such as those used in evidence-based medicine. However, a unique approach is needed to summarize a body of DDI evidence, which often consists of case reports, retrospective reviews, and extrapolation from in vitro studies, with few controlled clinical studies conducted in relevant populations [17–19]. Some DDIs do not require randomized controlled trials to confirm their existence. It is possible to reasonably extrapolate many interactions based on pharmacokinetic and/or pharmacodynamic properties of a drug without

placing patients at unnecessary or unethical risk. High-quality observational studies and evidence from real-world use can be applied to confirm the association with adverse clinical outcomes and to evaluate the magnitude of harm and relevant risk factors. Therefore, evidence supporting a DDI may be derived from what would be regarded as less rigorous study designs for other research questions.

### 3.2.1 Existing DDI Evidence Evaluation Methods

We conducted a nonsystematic search for published methods for evaluating DDI case reports using MEDLINE and also queried workgroup members for relevant articles and instruments (see ESM 3). Only one instrument, the

Drug Interaction Probability Scale (DIPS) [20], was identified to be specifically developed to evaluate individual case reports for DDIs. This 10-item scale was designed to assess an adverse event for causality by a DDI. DIPS was developed to address the limitations of previous assessment instruments, such as the Naranjo scale [28], that failed to evaluate causality associated with concomitant medications. DIPS takes into consideration previous credible reports, consistency with known interactive properties, time course of the interaction, results of de-challenge and re-challenge, and alternative explanations. DIPS also meets published criteria for assessing causality in terms of guiding users to conduct an explicit, transparent, complete, and balanced assessment of the attributes important to causality assessment of whether an adverse drug interaction occurred and exists [29].

We also searched for published literature of methods that evaluate a collection of evidence relevant to establishing that a DDI exists (see ESM 3) and found two systematic approaches [30, 31]. The first is the approach used for developing a DDI knowledgebase in Swedish and Finnish computerized CDS systems (SFINX) [30]. This system categorizes level of documentation (0–4) and clinical relevance (A–D). A ‘0’ level of documentation reflects potentially dangerous interactions that have not been, and probably never will be, documented in clinical studies. The second approach described in the literature is a systematic assessment of DDIs for CDS systems in the Netherlands [31]. Four core parameters are used to assess each DDI: evidence supporting the interaction; clinical relevance of the potential adverse reaction; risk factors related to patient, drug, or disease characteristics; and incidence of the adverse reaction. A 5-item scale is used to assess the quality of evidence for a DDI. The approach requires the existence of a reasonable mechanistic explanation in order to establish a DDI based solely on pharmacokinetic or pharmacodynamic properties.

While the DIPS approach to case report evaluation was considered acceptable, the existing methods for evaluating a body of evidence were considered more complex than necessary because they combine DDI evidence assessment with questions of clinical relevance. Additionally, existing methods do not explicitly address reasonable extrapolation of DDIs from *in vitro* findings, nor do they explicitly address study quality and interpretation in the context of DDIs.

### 3.2.2 *The Need for a New DDI Evidence Evaluation Instrument*

Given the limitations of the available tools, we agreed that a new assessment instrument was needed to objectively evaluate a body of evidence to establish the existence of a

DDI. It was further agreed that this instrument should include concepts from previously published DDI evidence rating instruments [30, 31] but with fewer categories based on the presence or absence of specific, clearly defined, types of evidence.

Specific guidance is needed for reasonably extrapolating DDIs that are unlikely to be evaluated in clinical trials. Reasonable extrapolation refers to using the knowledge of the mechanism of a DDI to predict the risk of a DDI from one pair of drugs to multiple pairs with similar pharmacologic properties. Extrapolation of pharmacodynamic interactions is commonplace. For example, not every possible combination of a benzodiazepine and ethanol has been studied. Yet, all benzodiazepines are assumed to interact in a similar manner with ethanol. DDIs based on pharmacokinetic mechanisms present more of a challenge to extrapolation because numerous elimination pathways may be involved and it is difficult to predict the magnitude of the interaction without additional data.

Based on the needs described above for a new method to establish the existence of a DDI, we developed the DRUG Interaction eVIDence Evaluation (DRIVE) instrument. The rationale behind the DRIVE instrument is to (i) use simple evidence categories; (ii) include causality assessment with DDI case reports (via DIPS); (iii) apply reasonable extrapolation, including from *in vitro* studies; (iv) address evidence/statements provided in product labeling; and (v) describe study quality criteria and interpretation in the context of DDIs. The purpose of the DRIVE instrument is to promote consistent, transparent, and systematic evaluation of a body of evidence to establish that a DDI exists. Once formally evaluated and validated, the DRIVE instrument may be used by drug compendia and knowledgebase editors who develop and maintain DDI content for drug information and decision support systems. Health professionals, researchers, and journal editors may also find the instrument useful. Systematic evidence review should include a thorough search for relevant published and unpublished literature and, therefore, we also recognized that future work should seek to develop systematic methods for conducting literature searches for assembling DDI evidence [32–36].

Due to regulatory and clinical practice implications, special attention was given to evidence from product labeling. Formal regulatory documents, such as product labeling, preapproval reviews, and post-marketing analyses are important sources when evaluating DDI evidence [37]. Preapproval reviews and post-marketing analyses by the sponsor and the FDA can identify unpublished DDI evidence as the documents are generally more thorough and may be useful when information in product labeling is incomplete [36]. Despite significant challenges and effort involved in reviewing FDA materials [37], a careful review



of these documents should be performed when conducting an evaluation of DDI evidence [36].

In our collective experience and opinion, the majority of pre-marketing DDI studies conducted and described in product labeling are well designed and executed. However, we agree that DDI content in drug compendia and CDS systems do not always need to align with product labeling, even when that information is listed as a contraindication or boxed warning. This opinion is based upon examples in which the labeling is not consistent with existing evidence [38] and significant variation occurs in the DDIs listed in product labeling compared with published information [33, 34, 39–41]. We acknowledge that the purpose and guidance of labeling is unique, and also that the FDA has taken important steps to improve the quality and usefulness of DDI information in product labeling [42–44]. Continued effort is recommended to improve the consistency and timeliness of DDI information in product labeling, particularly for older nonproprietary drugs.

### 3.2.3 Assessing Clinical Relevance

If there is sufficient evidence that a DDI might require clinical management, further evaluation is needed to establish the clinical relevance. Clinical outcomes associated with the DDI must be determined, including the magnitude, variability and frequency of effects (if known), and modifying factors that may increase or decrease the risk of patient harm. Depending on the context, exposure to a clinically relevant DDI might warrant a change in therapy, increased monitoring, and/or patient education. Assessing the clinical relevance of a DDI is an estimate, at best, because inter-patient variability is often unknown, and for pharmacokinetic DDIs, changes in the object drug can vary four- to six-fold [45, 46]. For some DDIs, it is reasonable to assign a general risk rating based on the properties of the object drug, such as those with a NTI.

We recommend providing estimates of the frequency (incidence) of adverse outcomes from DDIs when available. However, assessing these frequencies is difficult because data are largely limited to observational studies, which are susceptible to confounding. When available, the definition of the adverse outcome and the applicable population should be clearly specified. For example, patients can be informed that the risk of upper gastrointestinal (GI) bleeding is estimated to increase by 19 % with combined use of selective serotonin reuptake inhibitors (SSRIs) and nonsteroidal anti-inflammatory drugs (NSAIDs) beyond the effects of each individual drug [47]. Estimates can also be provided that 179 (95 % CI 107–319) high-risk patients (e.g., elderly, previous GI bleed) and 645 (95 % CI 387–1,152) low-risk patients need to be treated with the SSRI and NSAID combination to cause one upper-GI bleed

[47]. But for most DDIs—even those that are well documented and potentially dangerous—only rough estimates of the incidence of adverse outcomes are known.

Thorough evidence evaluation of DDI literature should include documented methods to mitigate harm (e.g., dosage adjustment, monitoring strategies, and therapeutic alternatives) [22, 48]. Reasonable therapeutic alternatives may include DDIs ruled out by mechanistic principles, preferably with one or more negative studies with limited bias and confounding.

We also recommend that modifying factors that may increase (risk factors) or decrease (mitigating factors) susceptibility to DDIs should be considered when evaluating and reporting DDI evidence [48]. Drug-related modifying factors may include dose, duration, route of administration, order of administration, timing of dose, and co-medications. Patient-related modifying factors may include age, sex, pharmacogenomics [49], comorbidity, clinical status, vital signs, laboratory values, and indication for the drug. Identifying modifying factors is essential because research shows that providing patient-specific risk factors in CDS improves the specificity of alerts [50, 51]. There are many situations where a particular DDI may not be clinically relevant to a specific patient due to mitigating factors that result in a negligible risk of adverse outcomes. For example, a precipitant drug is unlikely to produce a clinically relevant DDI for a patient with a genetic variant producing a nonfunctional cytochrome P450 (CYP) enzyme (i.e., poor metabolizer) [52]. However, information on factors that alter patient susceptibility to DDIs is not yet systematically reviewed in DDI guidelines [51]. In general, more research is needed to identify modifying factors to inform CDS algorithms and clinical decision making.

More work is also needed to identify the most appropriate process to rate the quality of DDI evidence and provide graded recommendations to reduce the risk of adverse consequences [48]. We recommend considering the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system, a well accepted standard to indicate quality of evidence and strength of recommendations [53–56]. For example, the University of Liverpool has adapted the GRADE approach for communicating recommendations related to drug interactions with medications used for HIV and hepatitis C (described in ESM 4) [57, 58].

### 3.3 Key Question 2: What Evidence is Required for a DDI to be Applicable to an Entire Class of Medications?

CDS systems can generate nuisance alerts when they inappropriately define or represent a DDI as a ‘class’ effect.

Knowledge of the mechanism of interaction is crucial to determining whether there is basis for a class effect. Most class-based DDIs are of a pharmacodynamic nature, with additive [e.g., angiotensin-converting enzyme (ACE) inhibitors + angiotensin II receptor blockers (ARBs)] or opposing (e.g.,  $\beta$ -blocker +  $\beta$ -agonist) pharmacologic effects. In contrast, pharmacokinetic interactions are rarely generalizable to all agents within a class [22, 59]. Even when there is seemingly a class effect, the magnitude of effect can vary, which often makes it necessary to consider each drug in the class individually. For example, azole antifungal agents can inhibit CYP3A4. However, itraconazole and ketoconazole are much more potent inhibitors than fluconazole, so the magnitude of interaction with drugs that are primarily metabolized by CYP3A4 may differ significantly, which would impact the clinical relevance of the interaction. This can be shown by their effect on triazolam levels: itraconazole and ketoconazole increase the AUC of triazolam by 27- and 22-fold, respectively [60], whereas fluconazole causes a 4.4-fold increase in AUC [61].

The overwhelming majority of extrapolated DDIs are pharmacodynamic, because few studies are conducted to investigate this type of interaction. In the absence of drug-specific data, a class-based interaction may be reasonably assumed if the purported mechanism of interaction is biologically plausible and consistent with known pharmacology of one or both classes of drugs involved. Class examples include SSRIs plus other serotonergic drugs related to serotonin syndrome, and anticoagulation plus antiplatelet agents related to bleeding.

Occasionally, pharmacokinetic interaction data may be extrapolated from one agent to other agents in the class if the purported mechanism of interaction involves common pharmacologic effects. For instance, NSAIDs may reduce the renal excretion of lithium and therefore increase the risk of toxicity [62]. The proposed mechanism of interaction is inhibition of renal prostaglandin synthesis by NSAIDs, which leads to reduced renal blood flow. Although lithium toxicity has not been reported with all NSAIDs, the interaction is probably applicable to the entire class based on their common ability to inhibit prostaglandin synthesis.

We recommend that DDIs should be class-based only when the evidence (or reasonable extrapolation) applies to the entire pharmacological class of drugs.

### 3.4 Key Question 3: How Should a Structured Evaluation Process Be Vetted and Validated?

In Key Question 1, we recommended use of a new instrument as a standard to evaluate DDI evidence. However, any new DDI evidence evaluation instrument should undergo a

rigorous evaluation. The evaluation should ensure that the instrument is easy to apply by end users and produces results that are generally concordant with other DDI evidence rating systems, except where differences are expected.

We recommend evaluating any new assessment tools using a subset of 15 ‘high-priority’ DDIs (drug pairs which should always generate an alert) approved by a panel of experts commissioned by the US ONC [63]. There are also several existing studies that have systematically collected evidence for a set of DDIs and then examined concordance on DDIs mentioned in drug information sources [12, 34, 39, 41, 64]. These studies can provide DDIs for which there are varying degrees of agreement across drug information sources (e.g., some DDIs that all sources mention and others that only one source mentions).

Evaluation tools should provide clear definitions and be easy to use. The results from such tools should be internally consistent as well as across users. Measures of agreement among experts, such as Kappa statistics, should be reported in validation studies.

## 4 Discussion

This expert workgroup was convened to recommend an approach for evaluating DDIs in order to provide consistent, evidence-based CDS systems for healthcare providers. Because of the numerous challenges to evaluating evidence for DDIs, we propose a systematic and transparent process to evaluate evidence that supports the existence of clinically relevant DDIs. Furthermore, the use of a standardized evidence-based approach to evaluate DDI evidence will eliminate combinations with a low probability of harm and minimize legal liability for knowledgebase vendors and healthcare systems [23].

Our search for relevant tools to evaluate the DDI evidence identified no instruments that possessed all of the attributes believed to be important. Consequently, DRIVE was developed for evaluating the body of evidence for DDIs using important concepts from existing evidence evaluation methods with a focus on simplicity and explicit criteria for certain types of evidence [30, 31]. In this process, several terms were defined for use when evaluating DDIs because consistent application of terminology is requisite for systematic evaluation.

We recommend that any systematic approach to evaluating DDI evidence be validated to ensure the method is worthwhile. To that end, further evaluation of the DRIVE approach, including explicit criteria for what constitutes a well designed and executed study should be developed. DDI evidence reported in product labeling should be evaluated by the same criteria as published studies to establish sufficiency of DDI evidence.

For case reports, DIPS was judged to be the most appropriate published method to evaluate whether a DDI occurred [20]. Case reports may provide the first evidence of DDIs; however, using these reports as the sole evidence source has several disadvantages. They are often poorly described, leading to speculation and potentially inaccurate inferences about causal relationships. Use of case reports that are later found to be incorrect results in erroneous data listed in prescription product labeling and/or drug information compendia that are very difficult to correct. Therefore, careful evaluation of case reports is needed to establish the existence of a DDI.

There are limitations that should be kept in mind when employing experts in consensus recommendations. The most significant limitation is that results from consensus groups are driven by the membership. To overcome this limitation we invited members from diverse backgrounds. We refer the reader to other reports that outline other limitations of consensus groups [65].

We accomplished our goal of identifying principles for establishing a systematic process for evaluating evidence for DDIs; however, more work remains in certain areas. Although the DRIVE instrument may be used in the future to affirm that a DDI exists, further evaluation is needed to establish criteria for assessing clinical relevance. This entails identifying the associated clinical effects and their magnitude, variability, and estimated frequency. Modifying factors that may increase or decrease the risk of patient harm should also be identified. Pharmacogenetic research can further improve the precision of DDI evidence and CDS by identifying patient-specific predisposing factors. More work is also needed to identify the most appropriate process to rate the quality of DDI evidence and provide graded recommendations to reduce the risk of adverse consequences [48]. We recommend considering the GRADE system, a well accepted standard to indicate quality of evidence and strength of recommendations [53–56].

## 5 Conclusion

We convened a group of experts in pharmacology, drug information, biomedical informatics, and CDS to develop recommendations for best practices for evaluating DDI evidence. To promote agreement among DDI information systems we recommend consistent use of relevant terminology, including DDI, potential DDI, clinically relevant DDI, and seriousness. We created a definition of what constitutes an NTI drug specifically related to evaluating the clinical relevance of a potential DDI. For evaluating case reports of DDIs, we recommend the use of the DIPS tool. When formally validated, we recommend the use of

the DRIVE instrument to evaluate a body of evidence for existence of a DDI. Defining the clinical relevance of a DDI is extremely important because of thousands of theoretical, but not clinically relevant, interactions. Evaluations of evidence should consider modifying factors as well as the frequency of occurrence (if known). Broad indictments for drugs in the same therapeutic class should be done cautiously and only when there is sufficient evidence based on biological plausibility and known pharmacology. Finally, we recommend that any tools used to assess DDI evidence undergo a rigorous validation to ensure the approach results in reproducible and consistent findings.

**Acknowledgments** This project was funded by the Agency for Healthcare Research and Quality (AHRQ) Grant No. 1R13HS021826-01 (Malone DC-PI). RDB received support to contribute to this project from AHRQ Grant No. K12 HS019461-01, National Library of Medicine Grant No. R01 LM011838-01, and National Institute of Aging Grant No. K01 AG044433-01. Additional support was provided by Cerner, Elsevier Clinical Solutions, Epocrates, athenahealth, Inc., FDB (First Databank, Inc.), Truven Health Analytics, and Wolters Kluwer. The authors would like to thank Alissa Rich for her participation in workgroup discussions and Loretta Peters for her editorial assistance in preparing this manuscript. We would like to acknowledge that John Horn developed the Drug Interaction Probability Scale that is recommended for use in evaluating drug interaction case reports.

**Conflicts of interest** Lisa E. Hines, Mary Brown, and Amy J. Grizzle have no conflicts of interest to report except for grant support and travel support from the University of Arizona. Richard T. Scheife, Richard D. Boyce, John Horn and Jeremiah D. Momper have no conflicts of interest to report except for travel support from the University of Arizona to attend workgroup meetings. Christine D. Sommer is an employee of FDB, which supported the travel to workgroup meetings and reviewed the manuscript during business hours. Stephen J. Sklar is an employee of Wolters Kluwer, which provided a portion of financial support for this project and provided support for his travel to meetings. Sophie P. Chung and Susan Comes are employees of Epocrates, athenahealth, Inc., which paid for their travel to workgroup meetings. Gretchen Jones was employed by Epocrates, athenahealth, Inc., which paid for the travel to meetings, and she holds stock in that company. Clarissa Borst is an employee of Elsevier, which provided support to the conference. Daniel C. Malone received support from the University of Arizona to conduct this project, which was partially funded by a conference grant from the Agency for Healthcare Research and Quality. Michael A. Wittie, Darrell R. Abernethy, Samantha K. Wong and Tricia Lee Wilkins have no conflicts of interest that are directly relevant to the content of this study.

## References

1. Aspden P, Institute of Medicine (U.S.). Committee on identifying and preventing medication errors. Preventing medication errors. Washington, DC: National Academies Press; 2007.
2. Magro L, Moretti U, Leone R. Epidemiology and characteristics of adverse drug reactions caused by drug–drug interactions. *Expert Opin Drug Saf.* 2012;11(1):83–94.
3. Centers for Disease Control and Prevention FASTSTATS—Emergency Department Visits. 2012. US Department of Health



- and Human Services. [cited 12/06/2013]; Available from: <http://www.cdc.gov/nchs/fastats/ervisits.htm>.
4. Centers for Disease Control and Prevention FASTSTATS—Hospital Utilization. 2010. US Department of Health and Human Services. [cited 12/06/2013]; Available from: <http://www.cdc.gov/nchs/fastats/hospital.htm>.
  5. Centers for Medicare and Medicaid Services. Eligible Professional Meaningful Use Core Measures: Measure 2 of 15. 2010. US Department of Health and Human Services. [cited 01/09/2013]; Available from: [http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/2\\_Drug\\_Interaction\\_ChecksEP.pdf](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/2_Drug_Interaction_ChecksEP.pdf).
  6. Hatton RC, Rosenberg AF, Morris CT, McKelvey RP, Lewis JR. Evaluation of contraindicated drug–drug interaction alerts in a hospital setting. *Ann Pharmacother*. 2011;45(3):297–308.
  7. Shah VS, Weber RJ, Nahata MC. Contradictions in contraindications for drug–drug interactions. *Ann Pharmacother*. 2011;45(3):409–11.
  8. Horn JR, Hansten PD. Comment: evaluation of contraindicated drug–drug interaction alerts in a hospital setting. *Ann Pharmacother*. 2011;45(6):826 (author reply-7).
  9. Kesselheim AS, Cresswell K, Phansalkar S, Bates DW, Sheikh A. Clinical decision support systems could be modified to reduce ‘alert fatigue’ while still minimizing the risk of litigation. *Health Aff (Millwood)*. 2011;30(12):2310–7.
  10. Saverno KR, Hines LE, Warholak TL, Grizzle AJ, Babits L, Clark C, et al. Ability of pharmacy clinical decision-support software to alert users about clinically important drug–drug interactions. *J Am Med Inform Assoc*. 2011;18(1):32–7.
  11. Metzger J, Welebob E, Bates DW, Lipsitz S, Classen DC. Mixed results in the safety performance of computerized physician order entry. *Health Aff (Millwood)*. 2010;29(4):655–63.
  12. Abarca J, Colon LR, Wang VS, Malone DC, Murphy JE, Armstrong EP. Evaluation of the performance of drug–drug interaction screening software in community and hospital pharmacies. *J Manag Care Pharm*. 2006;12(5):383–9.
  13. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc*. 2006;13(2):138–47.
  14. Miller AM, Boro MS, Korman NE, Davoren JB. Provider and pharmacist responses to warfarin drug–drug interaction alerts: a study of healthcare downstream of CPOE alerts. *J Am Med Inform Assoc*. 2011;18(Suppl 1):i45–50.
  15. Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians’ decisions to override computerized drug alerts in primary care. *Arch Intern Med*. 2003;163(21):2625–31.
  16. Grizzle AJ, Mahmood MH, Ko Y, Murphy JE, Armstrong EP, Skrepnek GH, et al. Reasons provided by prescribers when overriding drug–drug interaction alerts. *Am J Manag Care*. 2007. p. 573–8 (United States).
  17. Caccia S, Garattini S, Pasina L, Nobili A. Predicting the clinical relevance of drug interactions from pre-approval studies. *Drug Saf*. 2009;32(11):1017–39.
  18. Holbrook AM, Pereira JA, Labiris R, McDonald H, Douketis JD, Crowther M, et al. Systematic overview of warfarin and its drug and food interactions. *Arch Intern Med*. 2005;165(10):1095–106.
  19. Huang SM, Strong JM, Zhang L, Reynolds KS, Nallani S, Temple R, et al. New era in drug interaction evaluation: US Food and Drug Administration update on CYP enzymes, transporters, and the guidance process. *J Clin Pharmacol*. 2008;48(6):662–70.
  20. Horn JR, Hansten PD, Chan LN. Proposal for a new tool to evaluate drug interaction cases. *Ann Pharmacother*. 2007;41(4):674–80.
  21. Hines LE, Malone DC, Murphy JE. Recommendations for generating, evaluating, and implementing drug–drug interaction evidence. *Pharmacotherapy*. 2012;32(4):304–13.
  22. Hansten PD, Horn JR, Hazlet TK. ORCA: Operational Classification of drug interactions. *J Am Pharm Assoc*. 2001;41(2):161–5.
  23. Ridgeley MS, Greenberg MD. Too many alerts, too much liability: sorting through the malpractice implications of drug–drug interaction clinical decision support. *Saint Louis Univ J Health Law Policy*. 2012;5(257):257–96.
  24. Oates JA. Chapter 5. The science of drug therapy. In: Brunton LL, editor. Goodman & Gilman’s *The Pharmacological Basis of Therapeutics*. 11th ed. McGraw-Hill, Medical Publishing Division; 2006.
  25. Hines LE, Murphy JE. Potentially harmful drug–drug interactions in the elderly: a review. *Am J Geriatr Pharmacother*. 2011;9(6):364–77.
  26. Talbot JCC, Aronson JK, Stephens MDB. Stephens’ detection and evaluation of adverse drug reactions : principles and practice. 6th ed. Chichester: Wiley; 2011.
  27. Food and Drug Administration. Title 21—Food and Drugs Chapter I—Food and Drug Administration; US Department of Health and Human Services Subchapter D—Drugs for Human Use. 2013. [cited 7/28/14]; Available from: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcr/CFRSearch.cfm?fr=320.33>.
  28. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther*. 1981;30(2):239–45.
  29. Hutchinson TA, Lane DA. Assessing methods for causality assessment of suspected adverse drug reactions. *J Clin Epidemiol*. 1989;42(1):5–16.
  30. Böttiger Y, Laine K, Andersson ML, Korhonen T, Molin B, Ovesjö ML, et al. SFINX—a drug–drug interaction database designed for clinical decision support systems. *Eur J Clin Pharmacol*. 2009;65(6):627–33.
  31. van Roon EN, Flikweert S, le Comte M, Langendijk PN, Kwee-Zuiderwijk WJ, Smits P, et al. Clinical relevance of drug–drug interactions : a structured assessment procedure. *Drug Saf*. 2005;28(12):1131–9.
  32. Boyce R, Collins C, Horn J, Kalet I. Computing with evidence Part II: an evidential approach to predicting metabolic drug–drug interactions. *J Biomed Inform*. 2009;42(6):990–1003.
  33. Boyce RD, Handler SM, Karp JF, Hanlon JT. Age-related changes in antidepressant pharmacokinetics and potential drug–drug interactions: a comparison of evidence-based literature and package insert information. *Am J Geriatr Pharmacother*. 2012;10(2):139–50.
  34. Boyce RD, Collins C, Clayton M, Kloke J, Horn JR. Inhibitory metabolic drug interactions with newer psychotropic drugs: inclusion in package inserts and influences of concurrence in drug interaction screening software. *Ann Pharmacother*. 2012;46(10):1287–98.
  35. McDonagh MS, Peterson K, Balshem H, Helfand M. US Food and Drug Administration documents can provide unpublished evidence relevant to systematic reviews. *J Clin Epidemiol*. 2013;66(10):1071–81.
  36. Turner EH. How to access and process FDA drug approval packages for use in research. *BMJ*. 2013;347:f5992.
  37. O’Connor AB. The need for improved access to FDA reviews. *JAMA*. 2009;302(2):191–3.
  38. DiNicolantonio JJ, Serebruany VL. Challenging the FDA black box warning for high aspirin dose with ticagrelor in patients with diabetes. *Diabetes*. 2013;62(3):669–71.
  39. Anthony M, Romero K, Malone DC, Hines LE, Higgins L, Woosley RL. Warfarin interactions with substances listed in drug information compendia and in the FDA-approved label for warfarin sodium. *Clin Pharmacol Ther*. 2009;86(4):425–9.
  40. Chao SD, Maibach HI. Lack of drug interaction conformity in commonly used drug compendia for selected at-risk dermatologic drugs. *Am J Clin Dermatol*. 2005;6(2):105–11.

41. Hines LE, Ceron-Cabrera D, Romero K, Anthony M, Woosley RL, Armstrong EP, et al. Evaluation of warfarin drug interaction listings in US product information for warfarin and interacting drugs. *Clin Ther*. 2011;33(1):36–45.
42. Food and Drug Administration. Guidance for industry: drug interaction studies—study design, data analysis, and implications for dosing and labeling recommendations: draft guidance. 2012. US Department of Health and Human Services. [cited 12/1/2014]; Available from: <http://www.fda.gov/downloads/drugs/guidance/compliance/regulatoryinformation/guidances/ucm292362.pdf>.
43. Food and Drug Administration. Information for Healthcare Professionals on FDA's New Prescribing Information for Drugs. US Department of Health and Human Services. [cited 10/29/10]; Available from: <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/LawsActsandRules/ucm084189.htm>.
44. Food and Drug Administration. Indexing structured product labeling for human prescription drug and biologic products. Docket No. FDA-2010-N-0256. Federal Register. 2010;2010(75):33312–3.
45. Horn JR, Hansten PD. Predicting the magnitude of drug interactions: the final frontier. *Pharmacy Times*. 2006;72.
46. de Leon J, Spina E, Diaz FJ. Pharmacokinetic drug interaction studies must consider pharmacological heterogeneity, use of repeated dosing, and translation into a message understandable to practicing clinicians. *J Clin Psychopharmacol*. 2009;29(3):201–5.
47. Anglin R, Yuan Y, Moayyedi P, Tse F, Armstrong D, Leontiadis GI. Risk of upper gastrointestinal bleeding with selective serotonin reuptake inhibitors with or without concurrent nonsteroidal anti-inflammatory use: a systematic review and meta-analysis. *Am J Gastroenterol*. 2014;109(6):811–9.
48. Floor-Schreuderling A, Geerts AF, Aronson JK, Bouvy ML, Ferner RE, De Smet PA. Checklist for standardized reporting of drug–drug interaction management guidelines. *Eur J Clin Pharmacol*. 2014;70(3):313–8.
49. Verbeurgt P, Mamiya T, Oesterheld J. How common are drug and gene interactions? Prevalence in a sample of 1143 patients with CYP2C9, CYP2C19 and CYP2D6 genotyping. *Pharmacogenomics*. 2014;15(5):655–65.
50. Tamblyn R, Egale T, Buckeridge DL, Huang A, Hanley J, Reidel K, et al. The effectiveness of a new generation of computerized drug alerts in reducing the risk of injury from drug side effects: a cluster randomized trial. *J Am Med Inform Assoc*. 2012;19(4):635–43.
51. Seidling HM, Klein U, Schaier M, Czock D, Theile D, Pruszydlo MG, et al. What, if all alerts were specific—estimating the potential impact on drug interaction alert burden. *Int J Med Inform*. 2014;83(4):285–91.
52. Miura M, Tada H, Yasui-Furukori N, Uno T, Sugawara K, Tateishi T, et al. Effect of clarithromycin on the enantioselective disposition of lansoprazole in relation to CYP2C19 genotypes. *Chirality*. 2005;17(6):338–44.
53. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is “quality of evidence” and why is it important to clinicians? *BMJ*. 2008;336(7651):995–8.
54. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6.
55. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490 (2004-06-17 21:56:41).
56. Grade Working Group. [cited 7/28/14]; Available from: <http://www.gradeworkinggroup.org/index.htm>.
57. The University of Liverpool. HEP-Drug Interactions. [cited 12/1/14]; Available from: <http://www.hep-druginteractions.org/>.
58. The University of Liverpool. HIV-Drug Interactions. [cited 12/1/14]; Available from: <http://www.hiv-druginteractions.org>.
59. Horn JR, Hansten, Philip D. “Classy” drug interactions. *Pharmacy Times*. 2005. [cited 12/1/14]; Available at: <http://www.pharmacytimes.com/publications/issue/2005/2005-06/2005-06-9585>.
60. Varhe A, Olkkola KT, Neuvonen PJ. Oral triazolam is potentially hazardous to patients receiving systemic antimycotics ketoconazole or itraconazole. *Clin Pharmacol Ther*. 1994;56(6 Pt 1):601–7.
61. Varhe A, Olkkola KT, Neuvonen PJ. Effect of fluconazole dose on the extent of fluconazole-triazolam interaction. *Br J Clin Pharmacol*. 1996;42(4):465–70.
62. Hughes BM, Small RE, Brink D, McKenzie ND. The effect of flurbiprofen on steady-state plasma lithium levels. *Pharmacotherapy*. 1997;17(1):113–20.
63. Phansalkar S, Desai AA, Bell D, Yoshida E, Doole J, Czochanski M, et al. High-priority drug–drug interactions for use in electronic health records. *J Am Med Inform Assoc*. 2012;19(5):735–43.
64. Abarca J, Malone DC, Armstrong EP, Grizzle AJ, Hansten PD, Van Bergen RC, et al. Concordance of severity ratings provided in four drug interaction compendia. *J Am Pharm Assoc* (2003). 2004;44(2):136–41.
65. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess*. 1998;2(3):i–iv (1–88).