Clinical validation of clinical decision support systems for medication review: A scoping review

Birgit A. Damoiseaux-Volman | Stephanie Medlock |
Delanie M. van der Meulen | Jesse de Boer |
Johannes A. Romijn | Nathalie van der Velde |
Ameen Abu-Hanna

The aim of this scoping review is to summarize approaches and outcomes of clinical validation studies of clinical decision support systems (CDSSs) to support (part of) a medication review. A literature search was conducted in Embase and Medline. In total, 30 articles validating a CDSS were ultimately included. Most of the studies focused on detection of adverse drug events, potentially inappropriate medications and drug-related problems. We categorized the included articles in three groups: studies subjectively reviewing the clinical relevance of CDSS’s output (21/30 studies) resulting in a positive predictive value (PPV) for clinical relevance of 4–80%; studies determining the relationship between alerts and actual events (10/30 studies) resulting in a PPV for actual events of 5–80%; and studies comparing output of CDSSs to chart/medication reviews in the whole study population (10/30 studies) resulting in a sensitivity of 28–85% and specificity of 42–75%. We found heterogeneity in the methods used and in the outcome measures. The validation studies did not report the use of a published CDSS validation strategy. To improve the effectiveness and uptake of CDSSs supporting a medication review, future research would benefit from a more systematic and comprehensive validation strategy.

KEYWORDS
adverse drug events, clinical decision support systems, inappropriate prescriptions, validation studies

1 | INTRODUCTION

Drug-related problems (DRP), such as adverse drug reactions (ADRs) and adverse drug events (ADEs), are common. For example, approximately 10% of the patients experience an ADR during hospital stay and approximately 20% of patients experience an ADE in inpatients settings.1,2 ADRs are associated with prolonged hospital stay.3 About 70% of the ADEs occurring in older hospitalized patients and between 16–42% of the ADEs occurring in outpatients are preventable.2,4 Individualized medication reviews can prevent ADRs/ADEs. During medication reviews, prescribed medications are evaluated by reviewing their indication, interactions and appropriateness to optimize medication use and improve health outcomes.5

Clinical decision support systems (CDSSs) can support medication reviews and detect potentially inappropriate prescribing.6 CDSSs can, for example, identify patients in need of a medication review, identify ADRs/ADEs or offer advice on how medications should be monitored.
or changed. Effectiveness of previous CDSS interventions supporting (part of) a medication review for hospitalized patients varied. CDSSs were mostly effective on process-related outcomes, but not yet on patient-related outcomes.7,8

An intervention with a CDSS is complex and must address the needs and wishes of future users. The Two-Stream Model and the GUIDES checklist describe factors that contribute to the success of CDSSs.9,10 These factors include data quality, underlying clinical knowledge and the presentation and quality of the output/advice. To ensure that the output of a CDSS is accurate, of high quality and useful for clinical practice, the system should be tested and validated.11 According to the World Health Organization, validation includes proving and documenting that a system leads to the expected results.12 A medication-related CDSS validation strategy consists of four steps: (1) technical validation to check whether the CDSS functions as expected, (2) retrospective validation to review whether the output is clinically relevant, actionable and useful, (3) prospective validation before implementation in a real-life EHR to check whether the CDSS fits in the workflow (e.g., timing and frequency), and finally (4) post-implementation validation for continuous improvement.13,14 The validation process can prevent irrelevant output for the clinical setting, alert fatigue or low user acceptance. Validation of a CDSS is, therefore, key for successful implementation.13,14 To date, a systematic review was published in 2012 on the detection of ADEs, which found that the accuracy of ADE rules was reported in only 50% of the ADE detection system studies.15 The rule accuracy was often poor and varied due to multiple factors including differences in validation methods.15 However, this 2012 systematic review focused solely on ADE detection systems and did not focus on other ways of clinical decision support for medication review. Currently, an overview of published validation methods and outcomes, for the different CDSSs to support (part of) a medication review, is missing. A scoping review can be conducted to systematically identify literature, to clarify concepts or identify knowledge gaps.16

Therefore, we conducted a scoping review with the aim of creating an overview of validated CDSSs supporting (part of) a medication review. Our secondary aim was to summarize and assess the methods and outcomes of the validation studies.

2 | METHODS

A literature search (see Appendix 1 in the Supporting Information) combining the categories corresponding to the concepts of “CDSS”, “medication review”, and “validation” was conducted in Embase and Medline up to 9 August 2021. Search results were deduplicated in Endnote X9.17 The references of the included articles were checked to find additional relevant articles. The study protocol was published on our departmental website (https://kik.amc.nl/KIK/reports/tech_reports.html), and the completed PRISMA checklist for scoping reviews can be found in Appendix 2 in the Supporting Information.18

2.1 | Eligibility

Articles were screened based on title and abstract by one reviewer (D.M. or B.D.) using pre-specified eligibility criteria. Title and abstract screening was performed using the Rayyan research app.19 Studies were included if they described the clinical validation of a CDSS (with or without a prediction model) for medication review, for adults (≥18 years), and were published in English in a peer-reviewed journal. After title and abstract screening, the full text of the remaining articles were also screened using the eligibility criteria. The full-text screening was performed by two researchers (D.M. and B.D.), independently. The results were compared, discussed and, if necessary, a third researcher (S.M.) was consulted. CDSS was defined as “any computer program designed to help healthcare professionals to make clinical decisions”.20 The Pharmaceutical Care Network Europe defines medication review as “a structured evaluation of a patient’s medicines with the aim of optimizing medicines use and improving health outcomes. This entails detecting drug related problems and recommending interventions”.3 We included CDSSs supporting any part of the medication review process. CDSSs during prescribing (order-entry pop-ups) were excluded. For clinical validation, we included all studies assessing the quality of the output of the CDSS.

2.2 | Data extraction and risk of bias assessment

Data extraction and assessment of the risk of bias were conducted by two researchers (D.M. and B.D.), independently. The results were compared, discussed and, if necessary, a third researcher (S.M.) was consulted. A template for data extraction was developed by J.d.B. and B.D. We extracted study and CDSS characteristics, validation methods and all outcomes related to clinical validation. We used these validation methods and outcomes in a narrative synthesis. The outcomes of individual studies were rated on relevance and applicability on a scale of 1–3 by two researchers (D.M. and B.D.), independently. Furthermore, we assessed the compliance of the studies to the framework of Scheepers-Hoeks et al.14 The references of the included studies were screened for additional articles. Risk of bias was assessed using an adapted version of the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2).21 Originally, this tool was intended for risk of bias assessment in studies determining diagnostic accuracy against a reference standard. We adjusted the signalling questions in the QUADAS-2 for applicability to CDSS validation studies. We did not exclude studies with a high risk of bias in the narrative analysis.

3 | RESULTS

3.1 | Study selection

Figure 1 shows the PRISMA flow diagram for article selection. The search in Medline and Embase identified 2434 articles, and search in additional sources identified nine further articles. After screening and
eligibility assessment, 30 studies were included in this scoping review. In total, three of the included 30 articles were identified through additional sources (reference checking).

### 3.2 Characteristics

The included studies had been performed in 12 countries (Table 1): Belgium, Canada, Germany, Ireland, Israel, Korea, the Netherlands, Portugal, Spain, Sweden, Switzerland and the United States. Studies were published in the years 1998–2020. Of the 30 studies, 26 were conducted in or used data of an inpatient hospital setting, three of primary care, three of an outpatient (clinic), and two of a nursing home.

The included studies used different clinical knowledge as input for their CDSSs. The alerts were based on literature, guidelines, experts’ experience, prevalence and product guidelines. In two studies, the CDSS was based on machine learning algorithms, including statistical analysis and deep learning with neural networks, and in one study the CDSS used an algorithm with risk scores based on literature.22–24 Five of the 30 studies described the use of a tool for potentially inappropriate prescribing such as the Beers criteria, or STOPP/START criteria.25–29 Of the included studies, 11 gave support to detect ADEs, six for potentially inappropriate medications (PIMs), five for DRPs, four for ADRs, three for medication surveillance, two for medication errors and two for drug-related hazardous conditions (DRHCs). The output of the CDSSs were presented as a list, (paper or electronic) report, database with alerts, an extra line in pharmacy module, alerts printed in the pharmacy or as a risk score.

### 3.3 Validation methods

All included studies used real-life patient data for the validation of the CDSSs. The studies included between 33 and 17878 patients. Four studies did not report information on the number of patients or admissions.23,30–32 Twenty-one (70%) studies reported how many (type of) rules were included in the knowledge base of the CDSS, ranging from 5 to 7879 alert (types). As Table 2 shows, all studies reported a frequency of generated alerts either as number of alerts or as number of patients/admissions with an alert. Of the 30 articles, eight studies used retrospective data for the validation (step 2 of the validation strategy of Scheepers-Hoeks et al.), 17 studies used prospective data (step 3), and five studies used post-implementation data for continuous improvement (step 4). One study assessed the CDSS with a pre-defined outcome threshold and aimed for a sensitivity of ≥80%.33 We divided the included studies into Group A, B and/or C (Table 2). Group A assessed clinical validation by subjectively checking the clinical relevance of the alerts (21 studies). Group B checked the association of alerts/advice with actual occurrence of events (10 studies) and group C compared the CDSS’s output with a chart/medication review in the whole study population (10 studies).

**Group A Clinical relevance of CDSS’s output**

In 21 of 30 studies, a reviewer checked the output of the CDSS and assessed the clinical relevance through their clinical experience.22,23,25–29,31,32,34–45 The reviewers assessed the clinical relevance with or without checking the charts or contacting a physician/
| Authors                  | Study setting | Patients                  | CDSS description                                                                 | CDSS input                                                                 | CDSS output                                                                 |
|--------------------------|---------------|---------------------------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|
| Arvisais et al. (2015)   | Hospital, all (except four) wards | $n = 182$ (200 patient-days) Mean age = 84 Female = 36% | Computerized alert system (CAS), used by pharmacists | $n = 5$ geriatric explicit criteria by expert panel based on Beers criteria and prevalence | Alerts (on HTML page), daily generated by database |
| Azaz-Livshits et al. (1998) | Hospital, medical ward | $n = 150$ patients Mean age = NR Female = 50% | ADR detection tool | $n = 14$ ADR signals from laboratory data | List of alerts |
| Buckley et al. (2018)   | Hospital, ICU and general ward | $n = 634$ patients Mean age = 60 non-ICU, 59 ICU Female = 7% non-ICU, 42% ICU | DRHC triggers | $n = 20$ unique trigger alerts | Paper and electronic reports (webpage within hospital) |
| Cossette et al. (2019)  | Primary care | $n = 369$ patients Mean age = 77 Female = 71% | Computerized alert system (CAS) | $n = 11$ types of geriatric explicit criteria based on 2015 Beers and STOPP v2 criteria | List of patients with PIMs and alert reasons |
| Dalton et al. (2020)    | Hospital      | $n = 204$ patients Mean age = 77 Female = 49% | SENATOR software, PIM CDSS for physicians | $n = 111$ recommendations based on STOPP/START v2 | Report per patient on paper and per email |
| DiPoto et al. (2015)    | Three hospitals, ICU and general ward | $n = 623$ patients Mean age = 60 non-ICU, 59 ICU Female = 49% non-ICU, 39% ICU | Automated surveillance trigger alerts for DRHCs | $n = 93$ logic-based trigger rules | Alerts in centralized database |
| Domann et al. (2000)    | Hospital, medical ward | $n = 379$ patients Age = 51 Female = 35% | Computer-based ADR monitoring system | $n = 20$ automated laboratory signals | Daily list of alerts with patient name, date of event |
| Eppenga et al. (2012)   | Hospital      | $n = 619$ patients Mean age = 53 Female = 54% | Two medication surveillance CDSSs: (1) Centrasys (ISOFT) (2) Pharmaps advanced CDSS | $n = NR$ CDSS 1 = G-standard (Dutch drug database), CDSS 2 = G-standard + additional expert-based rules | CDSS 1: Alerts at one point in time CDSS 2: Generation of alerts once a day |
| Ferrández et al. (2017) | Hospital      | $n = 17 878$ patients Mean age = 55 Female = 49% | DRP alert in pharmacy warning system, used by pharmacists | $n = 414$ medications (7879 alerts), based on literature, selected by two pharmacists | Extra line in pharmacy module. Alerts include patient, dose and recommendations |
| Fritz et al. (2012)     | Hospital, two general internal wards | $n = 100$ patients Median age = 59 Female = 42% | Three medication surveillance CDSSs: (1) Pharmavista (2) DrugReax (3) TheraOpt | NR | Alerts |
|                         | Nursing home  | $n = 115$ patients        | PIP screening (polimedications) | $n = NR$, STOPP criteria | Automated alerts |
TABLE 1 (Continued)

| Authors                     | Study setting                                                                 | Patients | CDSS description                                                                 | CDSS input                                                                 | CDSS output                                                                 |
|-----------------------------|-------------------------------------------------------------------------------|----------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|
| Garcìa-Caballero et al. (2018) | Mean age = 79  
Female = 62%                                                 |          |                                                                                 |                                                                             |                                                                             |
| Hammar et al. (2015)        | Two geriatric clinics, three primary care units  
n = 254 patients  
Mean age = 84  
Female = 64% | Electronic expert support system (EES) for DRPs, used by physicians  
n = NR, DDIs based on Swedish databank |                                                                             | Paper-based reports with potential DRPs                                 |
| Hedna et al. (2019)         | Data from 110 outpatient clinics, 51 primary care units, 29 departments in three hospitals  
n = 745  
Mean age = 75  
Female = 57% | PHARAO, risk scores for ADE  
n = 9 (p)ADE algorithm, database of potential interactions |                                                                             | Low, intermediate and high risk with advice                                 |
| Hwang et al. (2008)         | Hospital, two ICUs, five general wards  
n = 598 patients  
Age = NR  
Female = NR | ADE monitor, integrated in HIS  
n = 46, based on previously studied CDSS in literature, modified by panel |                                                                             | List of alerts with report including medication, ADE, bed location          |
| Ibìáñez-Garcìa et al. (2019) | Hospital  
n = 25 449 admissions  
Age = NR  
Female = NR | ADE CDSS (HIGEA)  
n = 211 clinical rules based on bibliographic search, clinical practice, team |                                                                             | Real-time list of alerts                                                  |
| Jha et al. (1998)           | Hospital, 9 units  
n = 21 964 patient-days  
Age = NR  
Female = NR | ADE monitor  
n = 52 ADE detection rules based on literature, study team |                                                                             | List of alerts with report including name, bed, event, condition           |
| Jha et al. (2008)           | Hospital, 5 units  
n = 2407 patients  
Av. age = 74  
Female = 53% | Dynamic pharmacovigilance  
n = NR, rule base, using combination of laboratory, medication, patient demographic data |                                                                             | List of alerts                                                               |
| Levy et al. (1999)          | Hospital, medical ward  
n = 192 patients  
Age = ~75% > 60  
Female = 47% | ADR detection tool  
n = NR, ADR signals from laboratory data (same tool as Azaz-Livshits et al) |                                                                             | List of alerts                                                               |
| Miguel et al. (2013)        | Hospital  
n = 118 patients  
Mean age = 60  
Female = 41% | ADR detection tool, stand-alone  
(patient data needs to be filled in by hand)  
n = 10 ADRs, selection drugs and ADRs based on hospital formulary, INFARMED, book |                                                                             | Alerts with suggested ADRs and frequent ADR for prescribed drugs            |
| Peterson et al. (2014)      | Hospital, general medicine, orthopaedics, urology  
n = 179 patients  
Mean age = 72  
Female = NR | PIM review dashboard  
n = NR, triggers based on Beers and STOPP criteria, anticholinergic risk scale, formulary |                                                                             | List of patients with PIM(s) sorted by highest risk                      |
| Quintens et al. (2019)      | Hospital  
n = NR  
Age = 47–74  
Female = NR | Check of medication appropriateness (CMA)  
n = 78 clinical rules for PIMs, DRPs and ADEs based on literature and guidelines, selected by team |                                                                             | Once a day generation of list of alerts on a worklist                     |
| Raschke et al. (1998)       | Hospital, non-obstetrical patients  
n = 9306 admissions  
Age = NR | ADE system  
n = 37 ADE rules defined by authors |                                                                             | Alerts printed in pharmacy                                                |
| Authors                  | Study setting                          | Patients                        | CDSS description                  | CDSS input                                      | CDSS output                                      |
|-------------------------|----------------------------------------|----------------------------------|-----------------------------------|------------------------------------------------|------------------------------------------------|
| Rommers et al. (2011)   | Hospital, general internal ward        | \(n = NR\)                       | ADE alerting system (ADEAS),     | \(n = 121\) clinical expert-based rules       | Every morning, list of alerts                    |
|                         |                                        | Age = NR                         | pharmacists                       |                                                 |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| Rommers et al. (2013)   | Hospital, six wards                    | \(n = 931\) patients            | ADE alerting system (ADEAS),     | \(n = 121\) clinical expert-based rules       | List of alerts for patients with possible ADE   |
|                         |                                        | Age = NR                         | pharmacists                       |                                                 |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| Roten et al. (2010)     | Hospital, internal medicine and        | \(n = 501\) patients            | DRP screening tool               | \(n = 6\) DRP queries based on literature,    | List of patients with possible DRPs             |
|                         | geriatric wards                        | Age = NR                         |                                   | experience pharmacists, another hospital        |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| Schiff et al. (2017)    | Outpatients                            | \(n = NR\)                       | Medication errors, outlier       | \(n = 1706\) alerts, machine learning         | Alerts with short explanation                   |
|                         |                                        | Age = NR                         | detection screening (MedAware)    | algorithms identifying outliers                 |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| Silverman et al. (2004)| Hospital                               | \(n = NR\)                       | ADE detection system             | \(n = NR\), ADE detection rules (modified     | List of alerts                                  |
|                         |                                        | Age = NR                         |                                   | version of Jha et al.)                         |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| Segal et al. (2019)     | Hospital, internal medicine department | \(n = 3160\) patients           | Medication errors/ADE CDSS       | \(n = NR\), machine learning algorithms       | Alerts after change in clinical state patient  |
|                         |                                        | Age = NR                         | (MedAware)                        | identifying outliers                            |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| de Wit et al. (2015)    | Hospital, nursing home                 | \(n = 900\) patients            | Medication surveillance CDSS,    | \(n = 39\) clinical rules based on product    | Alerts                                         |
|                         |                                        | Age = NR                         | stand-alone, pharmacists         | info, known ADEs, prescribing mistakes         |                                                 |
|                         |                                        | Female = NR                      |                                   |                                                 |                                                 |
| de Wit et al. (2016)    | Hospital, geriatric ward               | \(n = 33\) patients             | Medication review CDSS, stand-    | \(n = 469\) clinical rules based on literature,| DRP alerts with advice to prevent ADE           |
|                         |                                        | Mean age = 83                    | alone, used by pharmacists       | guidelines, protocols, multidisciplinary team  |                                                 |
|                         |                                        | Female = 45%                     |                                   |                                                 |                                                 |

ADE, adverse drug event; ADR, adverse drug reaction; CDSS, clinical decision support system; DRHC, drug-related hazardous conditions; DRP, Drug-related problems; ICU, intensive care unit; NR, not reported; PIM, potentially inappropriate medication; PIP, potentially inappropriate prescribing.
| Authors                  | Group A/B/C and description methods                                                                 | Outcomes, relevance (R), applicability (A)                                                                 |
|-------------------------|------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| Arvisais et al. (2015)  | A Two junior pharmacists (independently) analysed alerts for clinical relevance                       | Clinical relevance = 75% (149 with ≥1 clinically relevant alert/200 patient days)                        |
| Azaz-Livshits et al. (1998) | B Signals evaluated by expert team                                                                      | 37% (78 signals related to ADR/212 signals)                                                             |
|                         | C Clinical pharmacologist reviewed charts for ADRs, evaluated by expert team                           | 29% (25 admissions ADR & signal/86 admissions with alert)                                                 |
|                         |                                                                                                     | Sn = 66% (25 CDSS/86 admissions by expert team)                                                           |
|                         |                                                                                                     | Sp = 49% (56/115 admissions)                                                                            |
| Buckley et al. (2018)   | B Pharmacist reviewed alerts and patient charts to determine causality for DRHCs/ADEs using validated tools | PPV (DRHCs) = 29% (249/870 alerts)                                                                         |
|                         |                                                                                                     | PPV (ADEs) = 5% (47/870)                                                                                    |
| Cossette et al. (2019)  | A Clinical relevance assessment by two pharmacists based on clinical experience                        | Clinical relevance: 41% (34/83 alerts) 42% (27/65 patients with ≥1 alert)                                |
| Dalton et al. (2020)    | A Independent analysis by pharmacist and physician (6-point scale)                                     | Clinical relevance = 74% (681 relevant/925 alerts)                                                       |
| DiPoto et al. (2015)    | A Pharmacists assessed alerts (clinically relevant: pharmacists proposed change)                      | Clinical relevance: 40% (ICU, 90/226 alerts), 45% (general ward, 235/525 alerts)                        |
|                         | B Pharmacist assessed causality trigger and adverse events (subgroup analysis of 161 triggers of 19 rules) | PPV (DRHCs) = 71% (115/161 triggers)                                                                    |
| Dormann et al. (2000)   | B Team of pharmacologist, clinician and pharmacists assessed alerts and patient chart for ADR (Naranjo score) | PPV (ADRs) = 13% (63/501 alerts)                                                                        |
|                         | C Compare ADR detected by CDSS with ADRs from spontaneous reporting                                    | Relative Sn = 74% (34 ADRs CDSS/46 [all] ADRs)                                                          |
|                         |                                                                                                     | Relative Sp = 75%                                                                                        |
| Eppenga et al. (2012)   | A Two pharmacists independently assessed alerts of two CDSSs for clinical relevance (of pharmacist would take action) | PPV (clinical relevance) CDSS 1 = 6% (150/2607 alerts), PPV (clinical relevance) CDSS 2 = 17% (384/2256 alerts) |
| Ferrández et al. (2017) | A Alerts reviewed for clinical relevance (action needed) by pharmacists                               | Clinical relevance = 20% (2808/13 833 alerts)                                                            |
|                         | C Compare DRPs detected by CDSS with pharmacist review                                                 | 79% (2808 DRPs by CDSS/3552 [all] DRPs)                                                                  |
| Fritz et al. (2012)     | A Alerts reviewed for clinical relevance (of pharmacists would take action) by pharmacists AND sensitivity to detect all 33 relevant alerts (identified by pharmacist while reviewing alerts) | PPV (clinical relevance) = 6%, 8%, 8% (3/53, 29/364, 25/328 alerts), respectively                       |
|                         |                                                                                                     | Sn = 9%, 88%, 76% (3/33, 29/33, 25/33 relevant alerts), respectively                                    |
| Garcia-Caballero et al. (2018) | A A physician and psychiatrist reviewed alerts                                                         | Relevance = 12% (140/1155 alerts)                                                                          |
| Authors                  | Group A/B/C and description methods                                      | Outcomes, relevance (R), applicability (A)                                      | R   |
|--------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------------------|-----|
| Hammar et al. (2015)35   | A  | A physician reviewed alerts for clinical relevance as part of medication review | Clinical relevance = 68% (502/740 alerts)                                      | +   |
| Hedna et al. (2019)24    | B  | For each alert, it was determined whether it was related to symptoms         | PPV = 0.20–0.25 [low risk: 150/776, intermediate risk: 93/460, high risk: 53/208 alerts] | ++  |
|                          | C  | Pharmacists extracted symptoms associated with medications, checked by second reviewer | Sn = 0.12–0.37 [high–low, patients’ symptom & alert/ patients’ symptom from review] Sp = 0.78–0.95 [low–high] NPV = 0.89–0.90 [high–low] | ++  |
| Hwang et al. (2008)49    | B  | Pharmacist reviewed alerts and charts for association alert with ADE         | PPV (ADEs) = 21% (148/718)                                                    | +++ |
|                          | C  | Pharmacists (checked by five other pharmacists) reviewed charts for patients without alert for ADEs | Sn = 79% (148/187 ADEs)                                                      | +++ |
| Ibáñez-Garcia et al. (2019)36 | A  | Pharmacist reviewed alerts, and advised physician                            | 51% (554 with advice/1086 alerts)                                             | +   |
| Jha et al. (1998)50      | B  | Reviewer analysed alerts, charts for ADE association (checked by physician)  | PPV (ADEs) = 17% (450/2620 alerts)                                             | ++  |
|                          | C  | Reviewers (blinded to CDSS) conducted ADE detection study                    | ADEs detected by CDSS = 45% (275/675 [all ADEs])                               | +++ |
| Jha et al. (2008)37     | A  | Reviewer analysed 52% alerts, and contacted physician if necessary          | Clinical relevance = 11% (30 with contact/266 alerts)                          | +   |
|                          | B  | Chart review to identify (potential) ADEs in a sample of patients            | PPV (ADEs) = 23%                                                               | ++  |
|                          |     | (checked by physician)                                                     | PPV (pADEs) = 15%                                                             |     |
| Levy et al. (1999)51    | B  | Analyses of signals                                                         | 18% (signals related to ADR (52/all signals [295])                            | +++ |
|                          | C  | Team reviewed charts for ADRs                                               | Sn = 62% (40 ADR admissions by tool/65 [all ADR admissions])                    |     |
|                          |     |                                                                             | Sp = 42% (79/135 admission without ADR)                                        |     |
| Miguel et al. (2013)52  | B  | ADRs detected by CDSS reviewed for true ADRs                               | PPV = 80% (65 true ADRs/81 all suggested ADRs)                                | ++  |
|                          | C  | Chart review and assessment by CDSS in population                           | 83% (10 ADR CDSS/12 ADRs in chart review)                                     | +   |
| Peterson et al. (2014)25 | A  | Pharmacist reviewed patients on dashboard and advised physician              | 12% (22 with intervention/179 patients)                                      | +   |
|                          |     |                                                                             | 6% (31 with interventions/485 alerts [PIMs])                                   |     |
| Quintens et al. (2019)30 | A  | Pharmacist checked alerts for appropriateness (clinical relevance = electronic note or phone call to physician) | Clinical relevance = 8% (3205 with action/39 481 alerts)                     | +   |
| Authors                          | Group A/B/C and description methods                                                                 | Outcomes, relevance (R), applicability (A)                                      | R   |
|---------------------------------|-------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|-----|
| Raschke et al. (1998)⁴⁸         | Pharmacist/radiology technicians evaluated alerts and advised physician                                  | Relevance = 71% (794 with advice/1116 alerts)                                   | +   |
| Rommers et al. (2011)³¹         | Hospital pharmacists reviewed true positive alerts for clinical relevance (= started intervention)       | Clinical relevance = 19% (14 with intervention/72 true positive alert)           | +   |
| Rommers et al. (2013)³⁹         | Pharmacists reviewed alerts, contacted and advised physician/nurse                                     | PPV (clinical relevance) = 8% (204 with advice/2650 alerts)                    | +   |
| Roten et al. (2010)³³           | Pharmacists conducted medication review (blinded to CDSS) to identify DRPs                             | 324 patients (65%) with alert                                                 | +++ |
| Schiff et al. (2017)²³          | Chart of patients with an alert were reviewed for accuracy and clinical validity                       | 126 alerts: Accuracy = 93% (based on data)                                       | +   |
| Silverman et al. (2004)³²       | Pharmacists reviewed alerts, and advised physician (3 x with different ADE rules)                    | Rule effectiveness (clinical relevance) = 5%, 6%, 13% (169/3117, 452/7390, 792/6136 alerts), respectively | +   |
| Segal et al. (2019)²²           | Biweekly interviews to manually review alerts                                                        | 315 alerts: Accuracy = 89% (no data issues)                                     | +   |
| de Wit et al. (2015)⁶⁰          | Pharmacists reviewed alerts for clinical relevance (= advised physician)                              | Efficiency (clinical relevance) = 4% (147/4065 alerts)                          | +   |
| de Wit et al. (2016)⁶¹          | Pharmacist and geriatrician independently checked DRPs by CDSS                                         | Clinical relevance = 12% (70/574 alerts)                                       | +   |
|                                |                                                                                                       | Sn = 72.9% (51 relevant alerts also classified as relevant by CDSS/70 relevant alerts) Sp = 98.6% (497 irrelevant alerts also classified as irrelevant by CDSS/504 irrelevant alerts) |     |
|                                | Geronto-pharmacology meeting discussed DRPs from a medication review (blinded to CDSS)              | 20% (44 DRPs CDSS/223 DRPs medication review) 28% (70 DRPs CDSS/249 [all DRPs]) | ++  |

ADE, adverse drug event; ADR, adverse drug reaction; CDSS, clinical decision support system; DRHC, drug-related hazardous conditions; DRP, drug-related problems; Group A, studying clinical relevance of CDSS's output; Group B, CDSS's output and actual occurrence of DRPs (patients with alert); Group C, CDSS's output and chart/medication review in whole population; PIM, potentially inappropriate medication; PPV, positive predictive value; Sn, sensitivity; Sp, specificity.
| Authors                | Outcomes, relevance (R), applicability (A) | Compliance framework\(^a\) step 2, 3, 4 and outcomes |
|-----------------------|-------------------------------------------|-----------------------------------------------------|
|                       | A                                         | 2 | 3 | 4 | Outcomes                              |
| Arvisais et al. (2015) | +++                                       |   |   |   | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = no                             |
| Azaz-Livshits et al. (1998) | ++                                       | X |   |   | PPV (clinical relevance) = no         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = yes                            |
| Buckley et al. (2018)  | +++                                       |   |   | X | PPV (clinical relevance) = no         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = yes                            |
| Cossette et al. (2019) | +++                                       |   |   | X | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = no                             |
| Dalton et al. (2020)   | +++                                       | X |   |   | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = no                             |
| DiPoto et al. (2015)   | +++                                       |   |   | X | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = yes                            |
| Dormann et al. (2000)  | +++                                       |   |   | X | PPV (clinical relevance) = no         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = yes                            |
| Eppenga et al. (2012)  | +++                                       |   |   | X | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = no                             |
| Ferrández et al. (2017) | +++                                       |   |   | X | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = yes                            |
| Fritz et al. (2012)    | +++                                       |   |   | X | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = yes                            |
| Garcia-Caballero et al. (2018) | +++                                       |   |   | X | PPV (clinical relevance) = yes         |
|                       |                                           |   |   |   | NPV = no                               |
|                       |                                           |   |   |   | Extra = no                             |
| Authors                  | Outcomes, relevance (R), applicability (A) | Compliance framework\(^{14}\) step 2, 3, 4 and outcomes | Outcomes |
|--------------------------|-------------------------------------------|----------------------------------------------------------|-----------|
| Hammar et al. (2015)\(^{25}\) | +++                                       | _ X _                                                  | PPV (clinical relevance) = yes NPV = no Extra = no |
| Hedna et al. (2019)\(^{24}\)       | ++                                       | X _ _                                                  | PPV (clinical relevance) = no NPV = yes Extra = yes |
| Hwang et al. (2008)\(^{40}\)     | ++                                       | X _ _                                                  | PPV (clinical relevance) = no NPV = no Extra = yes |
| Ibáñez-García et al. (2019)\(^{36}\) | +++                                      | _ X _                                                  | PPV (clinical relevance) = yes NPV = no Extra = no |
| Jha et al. (1998)\(^{50}\)       | ++                                       | _ X _                                                  | PPV (clinical relevance) = no NPV = no Extra = yes |
| Jha et al. (2008)\(^{27}\)       | +++                                      | _ X _                                                  | PPV (clinical relevance) = yes NPV = no Extra = yes |
| Levy et al. (1999)\(^{51}\)      | ++                                       | _ X _                                                  | PPV (clinical relevance) = no NPV = no Extra = yes |
| Miguel et al. (2013)\(^{52}\)    | ++                                       | X _ _                                                  | PPV (clinical relevance) = no NPV = no Extra = yes |
| Peterson et al. (2014)\(^{25}\)  | +++                                      | _ X _                                                  | PPV (clinical relevance) = yes NPV = no Extra = no |
| Quintens et al. (2019)\(^{20}\)  | +++                                      | X _ _                                                  | PPV (clinical relevance) = yes NPV = no Extra = no |
| Authors                      | Outcomes, relevance (R), applicability (A) | Compliance framework\(^{14}\) step 2, 3, 4 and outcomes                                      |
|-----------------------------|--------------------------------------------|------------------------------------------------------------------------------------------------|
|                             | A                                         | 2 3 4                                                                                             |
|                             |                                            | Outcomes                                                                                         |
| Raschke et al. (1998)\(^{38}\) | +++                                       | -- X --                                                                                          |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| Rommers et al. (2011)\(^{31}\) | +++                                       | -- X --                                                                                          |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| Rommers et al. (2013)\(^{29}\) | +++                                       | -- X --                                                                                          |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| Roten et al. (2010)\(^{33}\)  | +                                          | -- X --                                                                                          |
|                             |                                            | PPV (clinical relevance) = no                                                                    |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = yes                                                                                      |
| Schiff et al. (2017)\(^{23}\) | ++                                        | X --                                                                                              |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| Silverman et al. (2004)\(^{32}\) | +++                                       | -- -- X                                                                                          |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| Segal et al. (2019)\(^{22}\)  | ++                                        | -- X --                                                                                          |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| de Wit et al. (2015)\(^{40}\)  | +++                                       | -- -- X                                                                                          |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = no                                                                                       |
| de Wit et al. (2016)\(^{41}\)  | ++                                        | X --                                                                                              |
|                             |                                            | PPV (clinical relevance) = yes                                                                   |
|                             |                                            | NPV = no                                                                                         |
|                             |                                            | Extra = yes                                                                                      |

ADE, adverse drug event; ADR, adverse drug reaction; CDSS, clinical decision support system; DRHC, drug-related hazardous conditions; DRP, drug-related problems; Group A, studying clinical relevance of CDSS's output; Group B, CDSS's output and actual occurrence of DRPs (patients with alert); Group C, CDSS's output and chart/medication review in whole population; PIM, potentially inappropriate medication; PPV, positive predictive value; Sn, sensitivity; Sp, specificity.
nurse. In 13 of the 21 studies, the reviewer was a pharmacist(s), in two studies a combination of a pharmacist and a physician, in one study a combination of a pharmacists and a radiology technician, in two studies a physician and in three studies the profession of the reviewer was not described.

Group B CDSS’s output and occurrence of actual events (in patients with alert)

In 10 of 30 studies, a reviewer checked the patient charts to find a relationship between the output of the CDSS and actual occurrence of an ADR/ADE/DRHC. In eight of the 10 studies, all included patients with a CDSS alert were reviewed and in two studies, a sample of patients was reviewed.

Group C CDSS’s output and chart/medication review in the whole population (in patients with or without alert)

In 10 of 30 studies, the clinical validation was assessed by comparing the output of the CDSS with the output of a chart or medication review in the whole population (not only the patients with a CDSS alert). In four of the 10 studies, a pharmacist conducted a medication review to find DRPs/ADEs. In four other studies, a team or clinical pharmacist reviewed the charts for DRPs/ADEs/ADRs and two studies additionally used voluntary ADE/ADR reports. The reviewers were blinded to the output of the CDSS in four of the 10 studies.

3.4 Outcome measures and outcomes

Group A Clinical relevance of CDSS’s output

Table 2 shows that the 21 studies assessing the clinical relevance of CDSS’s output used different names to indicate positive predictive value (PPV) for clinical relevance: (clinical) relevance, clinical usefulness, clinical validity, efficiency, PPV and rule effectiveness. We used the term PPV (clinical relevance) for all studies.

Twenty studies reported the PPV (clinical relevance) of alerts, using a total of 89,974 alerts ranging from 53 to 39,481 alerts (median = 833 alerts, IQR = 318–2,639 alerts). A range of 4–80% of alerts were found to be clinically relevant (median = 13%, IQR = 8%–44%). One study reported the PPV (clinical relevance) in patient-days, finding that 75% of patient-days had ≥1 clinically relevant alert.

Group B CDSS’s output and occurrence of actual events (in patients with alerts)

All 10 studies reported the proportion of alerts related to a possible ADE/ADR/DRHC as an outcome. Eight of the 10 studies referred to this outcome as the PPV (Table 2). We used the term PPV for all studies. The PPV for ADEs varied between 5 and 23% (median = 19%, IQR = 14–22%). The PPV for ADRs varied between 13 and 80% (median = 28%, IQR = 17–48%). The PPVs for DRHCs were 29% and 71%.

Group C CDSS’s output and chart/medication review in the whole population (in patients with or without alert)

All 10 studies reported the sensitivity (Sn), or related measure, as the proportion of the DRPs/ADRs/ADEs found by CDSSs. The Sn for DRPs were 28%, 79% and 85%. The methods to detect ADRs varied and the reported outcomes were therefore not comparable: Sn = 62% (40/65 admissions, ADR detection by multiple methods), Sn = 66% (25/38 admissions, by expert team), Sn = 74% (34/46 ADRs, multiple methods), Sn = 83% (10/12 ADRs, by chart review). The Sn for ADEs were 45% and 79%. Of the 10 studies, five reported the specificity (Sp).

3.5 Risk of bias assessment

Most criteria of the assessment were categorized as low or unclear risk of bias for the included studies (Table 3). Twenty studies did not use a reference standard (chart/medication review for the whole population) and therefore the category was not applicable. Two studies scored a high risk of bias for one category.

4 DISCUSSION

We found 30 articles describing the clinical validation of a CDSS to support (part of) a medication review. Most of the studies focused on detection of ADEs, PIMs and DRPs. The methods used in these 30 articles can be categorized into three groups: (A) studies subjectively assessing the clinical relevance of CDSS’s output (21/30), (B) studies describing the association between alerts and actual events (10/30), and (C) studies comparing CDSS’s output against chart/medication review of the whole population (10/30).

In our scoping review, we found heterogeneity in the methods used and in the outcome measures. In alignment with our results, the systematic review from 2012 also found heterogeneity in the methods to assess the accuracy of ADE computerized detection rules. We, therefore, divided the studies into three groups. The studies in Groups A and B assessed the quality of the alerts (useful/correct) for patients with an alert. Studies in Group A assessed the quality of the alerts by clinical experience of reviewers. This is a relatively easily applicable method and effective as a first step in assessing the CDSS validity. Studies in Group B assessed the quality of the alerts through determining their association with medication errors (DRHCs, ADEs, ADRs). This increased the validation efforts of the reviewers but also increased the quality and relevance of the assessment. The studies in Group C assessed the quality of the alerts by studying patients with and without an alert (coverage).
| Authors                        | Risk of bias | Applicability concerns |
|-------------------------------|--------------|------------------------|
|                              | Patient selection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| Arvisais et al. (2015)        | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Azaz-Livshits et al. (1998)   | –            | ?          | ?                  | –              | –                | –          | –                  |
| Buckley et al. (2018)         | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Cossette et al. (2019)        | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Dalton et al. (2020)          | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| DiPoto et al. (2015)          | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Domann et al. (2000)          | –            | –          | –                  | –              | –                | –          | –                  |
| Eppenga et al. (2012)         | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Ferrández et al. (2017)       | –            | –          | –                  | –              | –                | –          | –                  |
| Fritz et al. (2012)           | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| García-Caballero et al. (2018) | ?             | –          | N/A                | N/A            | –                | –          | N/A                |
| Hammar et al. (2015)          | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Hedna et al. (2019)           | –            | –          | –                  | –              | –                | –          | –                  |
| Hwang et al. (2008)           | –            | –          | +                  | –              | –                | –          | –                  |
| Ibáñez-García et al. (2019)   | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Jha et al. (1998)             | –            | –          | –                  | –              | –                | –          | –                  |
| Jha et al. (2008)             | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Levy et al. (1999)            | –            | ?          | ?                  | –              | –                | –          | –                  |
| Miguel et al. (2013)          | –            | –          | –                  | –              | –                | –          | –                  |
| Peterson et al. (2014)        | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Quintens et al. (2019)        | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Raschke et al. (1998)         | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Rommers et al. (2011)         | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Rommers et al. (2013)         | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Roten et al. (2010)           | –            | –          | –                  | –              | –                | –          | –                  |
| Schiff et al. (2017)          | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| Silverman et al. (2004)       | ?            | –          | N/A                | N/A            | –                | –          | N/A                |
| Segal et al. (2019)           | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| de Wit et al. (2015)          | –            | –          | N/A                | N/A            | –                | –          | N/A                |
| de Wit et al. (2016)          | –            | –          | –                  | –              | –                | –          | –                  |

– Low, + High, ? Unclear.
The medication errors were found in the whole population through a medication review (or chart review, or spontaneous reporting) and these were compared with the output of the CDSSs. This is the most comprehensive and informative method as it gives insights into which medication errors the CDSS captured and which it missed. The outcome measures of Group A are included in the framework of Scheepers-Hoeks et al., but the reported methods and outcomes of Groups B and C are not (yet) fully included in the framework.14

The PPV (clinical relevance) of alerts reported in our review was 4–80% with a median of 13%. This median clinical relevance is low, when compared to the desired threshold of ≥89% mentioned in the framework of Scheepers-Hoeks et al.14 The PPV (clinical relevance) reported in the studies of our review were calculated using retrospective and prospective data, based on 53–39 481 generated alerts, 33–17 878 patients, 5–7879 number of rules, and different kinds of support (ADE detection, medication surveillance, PIMs, DRPs, DRHCs and medication errors). The two studies using machine learning as input (knowledge base) had a relatively high PPV (clinical relevance) of 75% and 80%. We did not find another variable or pattern that explained the difference between low vs high proportion of relevant alerts. The studies offered commentary on the reasons for low/high clinical relevance and for the impact of the clinical relevance. The clinical relevance would improve with higher data quality, by including more patient characteristics in the CDSSs and by the selection of high impact or prevalent problems.26,30,36,44,50 In the study by Dalton et al., a higher clinical relevance of the computerized recommendations was associated with an increase in implementation of the recommendations by the users.27 In the study by Segal et al., a new CDSS was compared to a legacy CDSS and this showed higher clinical relevance (85% vs 16%), lower alert burden (0.4% vs 37%) and more changes in practice (43% vs 5%) for the new CDSS.22 These outcomes are in line with a systematic review on factors influencing the uptake of medication-related CDSSs in hospitals. That review reported that the most frequent barriers were too many irrelevant alerts and information being not trustworthy.53 However, a lower clinical relevance of alerts may be accepted when an alert prevents severe adverse events.14

In 33% of the studies included in our review, the output of the CDSS was compared with a chart/medication review in the whole study population and therefore a CDSS coverage could be calculated (Group C). This percentage is comparable to the systematic review of Forster et al., in which nine out of 24 (38%) studies compared the accuracy of ADE detection rules to a chart review in the whole study population.15 The reported sensitivity (28–85%) and specificity (42–75%) in our review is within or near the broad range found by Forster et al. of 40–94% (Sn) and 1.4–89% (Sp).

Most studies in our scoping review (17/30) described step 3 (prospective validation) of the validation strategy of Scheepers-Hoeks et al.13,14 We reviewed clinical validation studies and therefore step 1 (technical CDSS validation) of the strategy was not within the scope of our review. None of the included studies reported all steps (2, 3 and 4) of the validation strategy and none referred to a published CDSS validation strategy. Furthermore, only one study described the expected or desired results (threshold) beforehand. Future studies would benefit from following a more structured and comprehensive validation strategy and including a pre-defined threshold for PPV, NPV, Sn or Sp.

A strength of our study is that we used a systematic approach with a broad search conducted in two databases and that the screening and data extraction was conducted by two researchers. Another strength is the use of a risk of bias assessment tool. We used an adapted version of the QUADAS-2 tool and two of the included studies had one domain flagged with a high risk of bias. Validation methods and outcomes not described in the studies were not included in the narrative analysis, which may have resulted in an underrepresentation of the used methods and outcomes.

In conclusion, 30 articles validating a CDSS to support (part of) a medication review were found. These studies varied markedly in methods and outcome measures. Most included studies described the clinical relevance of the alerts and only 33% described a comparison of CDSS’s output with all DRPs in the whole population. The studies did not report the use of a CDSS validation strategy. For the validation of their CDSSs, future studies would benefit from following a systematic and comprehensive approach as described in Scheepers-Hoeks et al., in order to limit irrelevant information, limit alert fatigue and, therefore, improve the uptake of CDSSs supporting (part of) a medication review.

ACKNOWLEDGEMENTS

The innovation funds of Amsterdam UMC, location AMC, supported this work. The sponsor had no role in the design, methods, data collection, analysis and preparation of this paper.

COMPETING INTERESTS

All authors have declared no conflicts of interest.

CONTRIBUTORS

B.D. was involved in study conception/design, data collection, analysis/interpretation of data, and drafting and revision of the manuscript. D.M. and J.B. were involved in study conception/design, data collection, analysis/interpretation of data and revision of the manuscript. J.R., A.A., S.M. and N.V. were involved in study conception/design, analysis/interpretation of data and revision of the manuscript. All authors approved the final version of the manuscript.

ORCID

Birgit A. Damoiseaux-Volman https://orcid.org/0000-0003-2779-8211
Stephanie Medlock https://orcid.org/0000-0002-2679-8095
Delanie M. van der Meulen https://orcid.org/0000-0002-1678-2057
Jesse de Boer https://orcid.org/0000-0001-9859-1774
Johannes A. Romijn https://orcid.org/0000-0003-0676-0409
Nathalie van der Velde https://orcid.org/0000-0002-6477-6209
Ameen Abu-Hanna https://orcid.org/0000-0003-4324-7954
medication errors in 100 medical inpatients. Eur J Clin Pharmacol. 2012;68(8):1209-1219. https://doi.org/10.1007/s00228-012-1241-6

35. Hammar T, Lidström B, Petersson G, Gustafson Y, Eiermann B. Potential drug-related problems detected by electronic expert support system: physicians’ views on clinical relevance. Int J Clin Pharmacol. 2015;37(5):941-948. https://doi.org/10.1007/s11096-015-0146-8

36. Ibáñez-Garcia S, Rodríguez-Gonzalez C, Escudero-Vilaplana V, et al. Development and evaluation of a clinical decision support system to improve medication safety. Appl Clin Inform. 2019;10(3):513-520. https://doi.org/10.1055/s-0039-1693426

37. Jha AK, Laguette J, Seger A, Bates DW. Can surveillance systems identify and avert adverse drug events? A prospective evaluation of a commercial application. J Am Med Inform Assoc. 2008;15(5):647-653. https://doi.org/10.1197/jamia.M2634

38. Raschke RA, Gollihare B, Wunderlich TA, et al. A computer alert system to prevent injury from adverse drug events. JAMA. 1998;280(15):1317-1320. https://doi.org/10.1001/jama.280.15.1317

39. Rommers MK, Zwaveling J, Guchelaar H-J, Teepe-Twiss IM. Evaluation of rule effectiveness and positive predictive value of clinical rules in a Dutch clinical decision support system in daily hospital pharmacy practice. Artif Intell Med. 2013;59(1):15-21. https://doi.org/10.1016/j.artmed.2013.04.001

40. de Wit HAJM, Mestres Gonzalvo C, Cardenas J, et al. Evaluation of clinical rules in a standalone pharmacy based clinical decision support system for hospitalized and nursing home patients. Int J Med Inform. 2015;84(6):396-405. https://doi.org/10.1016/j.ijmedinf.2015.02.004

41. de Wit HAJM, Hurkens KPGM, Mestres Gonzalvo C, et al. The support of medication reviews in hospitalised patients using a clinical decision support system. Springerplus. 2016;5(1):871. https://doi.org/10.1186/s40244-016-2376-1

42. Hulse RK, Clark SJ, Jackson JC, Warner HR, Gardner RM. Computer concepts. Am Hosp Pharm. 1976;33(10):1061-1064.

43. DiPoto JP, Buckley MS, Kane-Gill SL. Evaluation of an automated surveillance system using trigger alerts to prevent adverse drug events in the intensive care unit and general ward. Drug Saf. 2015;38(3):311-317. https://doi.org/10.1007/s40264-015-0272-1

44. Eppenga WL, Derijks HJ, Conenmans JMH, Hermens WAJJ, Wensing M, De Smet PAGM. Comparison of a basic and an advanced pharmacotherapy-related clinical decision support system in a hospital care setting in the Netherlands. J Am Med Inform Assoc. 2012;19(1):56-71. https://doi.org/10.1136/amiajnl-2011-000360

45. Ferrández O, Urbina O, Grau S, et al. Computerized pharmacy surveillance and alert system for drug-related problems. J Clin Pharm Ther. 2017;42(2):201-208. https://doi.org/10.1111/jcpt.12495

46. Azaz-Livshits T, Levy M, Sadan B, Shalit M, Geisslinger G, Brune K. Computerized surveillance of adverse drug reactions in hospital: pilot study. Br J Clin Pharmacol. 1998;45(3):309-314. https://doi.org/10.1046/j.1365-2125.1998.00685.x

47. Buckley MS, Rasmussen JR, Bikin DS, et al. Trigger alerts associated with laboratory abnormalities on identifying potentially preventable adverse drug events in the intensive care unit and general ward. Ther Adv Drug Saf. 2018;9(4):207-217. https://doi.org/10.1177/2042098618760995

48. Dormann H, Muth-Selbach U, Krebs S, et al. Incidence and costs of adverse drug reactions during hospitalisation. Drug Saf. 2000;22(2):161-169. https://doi.org/10.2165/00002018-200022020-00007

49. Hwang S-H, Lee S, Koo H-K, Kim Y. Evaluation of a computer-based adverse-drug-event monitor. Am J Heal Pharm. 2008;65(23):2265-2272. https://doi.org/10.2146/ajhp080122

50. Jha AK, Kuperman GJ, Teich JM, et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. J Am Med Inform Assoc. 1998;5(3):305-314. https://doi.org/10.1136/jamia.1998.0050305

51. Levy M, Azaz-Livshits T, Sadan B, Shalit M, Geisslinger G, Brune K. Computerized surveillance of adverse drug reactions in hospital: implementation. Eur J Clin Pharmacol. 1999;54(11):887-892. https://doi.org/10.1007/s002280050571

52. Miguel AI, Azevedo LF, Silva BM, Pereira AC. Resource-sparing computerized tool for detection of adverse drug reactions. Int J Pharm Technol. 2013;5(1):5118-5128.

53. Van Dort BA, Zheng WY, Baysari MT. Prescriber perceptions of medication-related computerized decision support systems in hospitals: a synthesis of qualitative research. Int J Med Inform. 2019;129:285-295. https://doi.org/10.1016/j.ijmedin.2019.06.024

SUPPORTING INFORMATION
Additional supporting information may be found in the online version of the article at the publisher’s website.

How to cite this article: Damoiseaux-Volman BA, Medlock S, van der Meulen DM, et al. Clinical validation of clinical decision support systems for medication review: A scoping review. Br J Clin Pharmacol. 2022;88(5):2035-2051. doi:10.1111/bcp.15160