Received: 30 May 2020 | Revised: 23 July 2020 | Accepted: 28 July 2020

DOI: 10.1111/bcp.14518

COMMENTARY

"The computer says no" Are there tools and algorithms that will help us stop potentially inappropriate medications?

Yoon Kong Loke | Katharina Mattishent

Norwich Medical School, University of East Anglia, UK

Correspondence
Professor Yoon K. Loke, Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ, UK.
Email: y.loke@uea.ac.uk

Emergency admissions due to adverse drug reactions (ADRs) have increased markedly in England from 60 000 hospitalizations per year in 2008 to 92 000 hospitalizations in 2015.[1] These ADRs are estimated to cost £83.7 million per year, with a projected loss of 627 lives.[2] More broadly, a meta-analysis of 42 studies from across the world found that 8.7% of hospital admissions of older people were due to ADRs.[3] In order to stem this tide, experts have emphasized the importance of de-prescribing high-risk drugs in primary care to "reduce appreciably the number of preventable drug-related admissions to hospital."[4] However, Elliott et al.'s evidence synthesis has reported that some ADR admissions cannot be avoided.[2] Resource-intensive, mass pre-emptive screening may only be worthwhile if a substantial proportion of hospitalized patients actually have preventable drug-related admissions.

A wide range of tools or algorithms have therefore been developed to enable more precise prediction of ADR risk, and targeted identification of susceptible patients for medication review. Typically, the tools are in the form of different sets of predictor variables or markers (derived through expert consensus) for screening and surveillance of patients with potentially elevated risk of ADRs. Some ADR prediction tools can be used to summarize an overall risk score for an individual based on patient characteristics, laboratory findings and the entire drug history. Alternatively, many tools take the approach of assessing ADR risk related to specific drug classes in a clinical condition that renders the patient particularly susceptible. Examples of specific markers are aspirin use in patients with history of peptic ulcer (where the patient is at risk of ADR of recurrent ulcer and bleeding), or beta-blockers in patient with asthma (where patient is at risk of ADR of worsening airway obstruction).[5]

These surveillance sets of ADR predictor markers can be deployed in different ways through General Practitioner's (GP) computer systems. Real-time alerts to warn of potential ADRs can be programmed (e.g. through clinical decision support systems) to pop up on the screen when GPs are seeing an individual patient who happens to trigger certain risk criteria. Conversely, a more formal approach, akin to surveillance or screening programmes, involves regular interrogation of the GP computer system for whole cohorts of patients who all share specific markers of ADR risk. Clinicians can run an electronic search of GP records, and then call back all those judged "at risk of ADR" for targeted medication review and de-prescribing to reduce subsequent harm.

Although both the above approaches seem eminently plausible, there are major issues with delivery and implementation into clinical practice. First is the (yet unproven) assumption that the chosen predictor variables are able to correctly and reliably distinguish cases who subsequently suffer ADRs from those who are unlikely to be harmed by ADRs. Here, we emphasize that the concept of "potentially inappropriate or hazardous medication" does not mean *definitely* inappropriate, nor that a harmful error has clearly occurred and must be rectified. Elliott et al.'s review estimated that only 1.7% of the prescribing errors in primary care would lead to serious harm, whereas the vast majority (98.3%) would have mild to moderate consequences for the health service.[2] We acknowledge that this classification is from a healthcare provider perspective and a patient may interpret the consequences to a different degree. The researchers concluded that "Very few, or no data were found that indicated direct links between errors and harm, … and what proportion of those errors reaching patients caused actual harm." Similarly, a meta-analysis of 67 studies found only a modest association (relative risk 1.25, 95% confidence interval 1.09–1.44) between "potentially inappropriate medications" (as measured by tools such as Beers or STOPP) and subsequent hospitalization with ADR.[6]

Moreover, current pop-up alerts and predictive ADR markers are often blunt tools that simply flag up crude binary categories of "at-risk" or "not at-risk". The absence of any actual numerical estimate of harm is unhelpful. Both clinicians and patients need data regarding the specific ADR risk (whether it be say, 4%, or 10%, or 50%) to make informed treatment choices, especially when trying to balance benefit of the drug against the harm. Clinicians may find it difficult to prioritize the highest risk individuals for urgent medication review because

existing tools may not provide sufficient depth for personalized risk stratification amongst "at risk" patients.

We also recognize that the vast bulk of research on predictor variables has been based on studies of serious ADRs in hospitalized patients and are unlikely to be fully applicable to primary care. The ADR predictor variables and tools derived from serious ADRs in hospitalized (tip of the iceberg) populations may not be generalizable to the wider community. There may exist a sizeable proportion of patients in primary care who suffer no serious harm from their medication (Figure 1). A computerized ADR prediction algorithm that is not well-validated in primary care could erroneously flag up large chunks of low-risk patients in the lower sections of Figure 1 because of low specificity and lack of understanding of clinical context. This, for instance, may include prevalent users where the GP has made a clinical assessment that the particular drug is actually not risky for the individual.

Prior to the deployment of any ADR tool, we must determine if the predictive markers of serious harm can or cannot perform accurately in a primary care setting. However, the available data on performance of ADR prediction tools is not reassuring. A systematic review of four ADR risk tools found that the discriminant ability was typically limited.[7] More recently, similarly disappointing results were obtained from evaluation of two different ADR prediction tools (including a newly derived tool incorporating STOPP/START items).[8] False positives and overestimation of ADR harm are the most important problems for de-prescribing. Healthcare professionals may end up being very frustrated with poorly specific tools that generate a lot of additional burdens and false alarms, whilst correctly predicting only small proportions of genuine ADRs. Patients can suffer unnecessary worry and inconvenience if they have been wrongly identified as being at risk of ADR. Moreover, specific predictors that overestimate risk of ADR may cause harm because beneficial drugs that are genuinely helping the patient are erroneously deprescribed.

Efficient use of scarce resources is also a key consideration, and a recently published model based on meta-analysis of primary care studies gives us some important insights. If screening and de-prescribing of inappropriate medicines succeeds in halving the number of inappropriate medications, between three or four ADR hospitalizations will be prevented for every 100 patients screened.[6] This translates to a number needed to be screened of between 25 to 33. Moreover, if each medication review requires 15 minutes of a GP or pharmacist's time, then they would have to set aside at least 8 hours of dedicated time in their work schedule specifically for medication review tasks to prevent one ADR hospitalization.

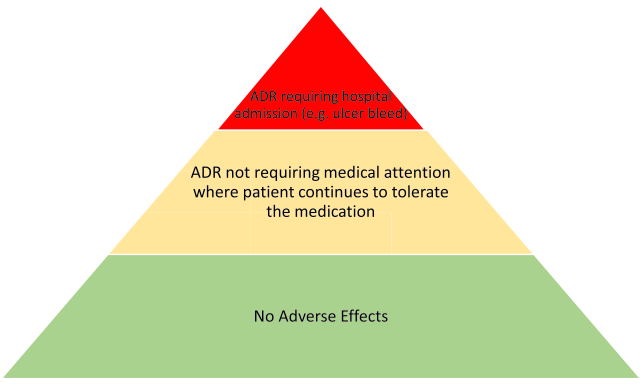

FIGURE 1 Hypothetical depiction of range of potential harm during follow-up of primary care patients that have been flagged up to be at risk of ADRs

The preceding paragraphs clearly demonstrate that ADR prediction tools and algorithms cannot be considered as universally effective panaceas for tackling hazardous prescribing and ADRs. But what should we be doing instead? We believe that the first step is to head back to the drawing board. The surveillance approach should be designed and tested in the same rigorous manner as any other screening procedure. Predictor variables for the ADR need to be obtained directly from the population of interest, e.g. primary care. We should ensure that these variables are available and reliably measured on the computer system where the tool or algorithm is deployed. Regularly updated and valid clinical information is particularly important for tools that utilize real-time alerts.

Once the predictor variables are assembled, the critical step is to then establish the natural history of those who are predicted to be "at-risk". Does the tool correctly flag up the cohort of patients who suffer serious ADRs if no additional preventive action is taken? What proportions are true-positive and false-positive cases of ADR during follow-up? Here, obtaining an estimate of the positive predictive value (PPV) is crucial to guiding clinical judgements on de-prescribing. In statistical terms, the PPV would be the probability of harm from ADR in those whom the screening test flagged up as at risk, i.e. of the people predicted by the screening tool/algorithm to be "at risk" of ADR, what is the likelihood that they will genuinely be harmed? Explicitly quantifying the performance of specific ADR markers in identifying "actual" harm will help clinicians and patients prioritize their efforts towards the areas of greatest risk, and to avoid situations where there a lot of false alarms.

Once the prognostic performance of the tool has been validated, the next step is to develop and test de-prescribing interventions targeted at high-risk medications responsible for serious ADRs. The intervention should be evaluated in pragmatic randomized controlled trials to confirm feasibility, as well as effectiveness on hard patient outcomes, rather than just number of prescriptions. There must be a clear specification for outcome measurement of collateral damage from the intervention which should cover both psychological and physical harms, and resource use. Potential adverse effects of screening and de-prescribing should include capture of loss of disease control from stopping a beneficial drug, as well as adverse reactions from switching or adding other drugs from the medication review (e.g. diarrhoea from adding proton pump inhibitor to someone taking dual antiplatelet therapy).

In summary, we feel that the time is not yet ready for computerized tools and algorithms to be used for de-prescribing and reducing ADRs. This is because the existing ADR prediction tools and algorithms do not have ability to synthesize clinical data to accurately estimate benefit–harm in individual patients. Although we strongly urge

physicians to keep a constant daily focus on de-prescribing, we cannot recommend implementation of any particular ADR prevention tool until we have robust and comprehensive studies that meet essential prerequisites for effective screening programmes.

Yes, it's good to regularly review medications and stop drugs … but this should be carefully targeted at the scenarios where we are sure that the harms outweigh benefits, as well as cases where there is no tangible evidence of any benefit.

COMPETING INTERESTS

There are no competing interests to declare.

ORCID

Yoon Kong Loke https://orcid.org/0000-0001-9109-2307
Katharina Mattishent https://orcid.org/0000-0002-7577-0689

REFERENCES

1. Veeren JC, Weiss M. Trends in emergency hospital admissions in England due to adverse drug reactions: 2008–2015. *J Pharm Health Serv Res*. 2017;8(1):5-11.
2. Elliott RA, Camacho E, Campbell F, et al. Prevalence and Economic Burden of Medication Errors in The NHS in England: Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU); 2018.
3. Oscanoa TJ, Lizaraso F, Carvajal A. Hospital admissions due to adverse drug reactions in the elderly. A meta-analysis. *Eur J Clin Pharmacol*. 2017;73(6):759-770.
4. Howard RL, Avery AJ, Slavenburg S, et al. Which drugs cause preventable admissions to hospital? A systematic review. *Br J Clin Pharmacol*. 2007;63(2):136-147.
5. Akbarov A, Kontopantelis E, Sperrin M, et al. Primary care medication safety surveillance with integrated primary and secondary care electronic health records: a cross-sectional study. *Drug Saf*. 2015;38(7):671-682.
6. Liew TM, Lee CS, Goh SKL, Chang ZY. The prevalence and impact of potentially inappropriate prescribing among older persons in primary care settings: multilevel meta-analysis. *Age Ageing*. 2020;49(4):570-579.
7. Stevenson JM, Williams JL, Burnham TG, et al. Predicting adverse drug reactions in older adults; a systematic review of the risk prediction models. *Clin Interv Aging*. 2014;9:1581-1593.
8. O'Mahony D, O'Connor M, Eustace J, Byrne S, Petrovic M, Gallagher P. The adverse drug reaction risk in older persons (ADRROP) prediction scale: derivation and prospective validation of an ADR risk assessment tool in older multi-morbid patients. *Eur Geriatr Med*. 2018;9:191-199.

How to cite this article: Loke YK, Mattishent K. "The computer says no" Are there tools and algorithms that will help us stop potentially inappropriate medications? *Br J Clin Pharmacol*. 2021;87:90–92. https://doi.org/10.1111/bcp.14518