ELSEVIER

feature

Development of predictive genetic tests for improving the safety of new medicines: the utilization of routinely collected electronic health records

Kevin Wing[1,*], kevin.wing@lshtm.ac.uk, **Ian Douglas[1]**, **Krishnan Bhaskaran[1]**, **Olaf H. Klungel[2,6]**, **Robert F. Reynolds[3]**, **Munir Pirmohamed[4]**, **Liam Smeeth[1]** and **Tjeerd P. van Staa[1,2,5]**

Serious adverse drug reactions are an important cause of hospitalization and can result in the withdrawal of licensed drugs. Genetic variation has been shown to influence adverse drug reaction susceptibility, and predictive genetic tests have been developed for a limited number of adverse drug reactions. The identification of patients with adverse drug reactions, obtaining samples for genetic analysis and rigorous evaluation of clinical test effectiveness represent significant challenges to predictive genetic test development. Using the example of serious drug-induced liver injury, we illustrate how a database of routinely collected electronic health records (EHRs) could be used to overcome these barriers by facilitating rapid recruitment to genome-wide association studies and supporting efficient randomized controlled trials of predictive genetic test effectiveness.

Adverse drug reactions are estimated to be responsible for over 5% of hospital admissions [1,2]. Approximately 150 drugs have been withdrawn from the market since 1960 owing to safety issues [3], in some cases many years after initial approval [4]. Evidence is increasing for a genetic predisposition to several serious adverse drug events [5]. For a limited number of drugs, observational genetic studies of association and subsequent randomized controlled trials (RCTs) of effectiveness of genotype-guided treatment have enabled predictive genetic tests to be developed that have allowed valuable medicines to remain on the market with greatly improved risk:benefit profiles [6].

Serious drug-induced liver injury is a leading cause of drug withdrawals [3]. Associations between specific genes and susceptibility to drug-induced liver injury caused by several drug therapies have been identified, although predictive genetic test development has been minimal and many genes conferring susceptibility are yet to be identified [7,8]. A major challenge is the time and cost associated with finding patients who have had a reaction of interest, and recruiting sufficient numbers for initial genome-wide association studies (GWAS) and subsequent replication studies [9]. Following predictive genetic test development, an RCT to evaluate effectiveness of the genotype-guided treatment versus standard care might be required [10], introducing further logistical challenges. The possibility of using databases of routinely collected EHRs to support pharmacogenomics has been discussed elsewhere (Yasmina *et al.* unpublished; [11–13]). In this article, we provide further detail by illustrating how an EHR database could be used to: (i) identify patients who have had serious adverse reactions linked to a newly licensed drug to invite them to provide genetic samples for GWAS; and (ii) test the efficacy of any developed genetic test in a cluster RCT. We illustrate these ideas using drug-induced liver injury as an example of an adverse drug reaction.

Identification of drug-induced liver injury within routinely collected EHRs

Routinely collected EHRs provide the potential for low-cost, efficient epidemiological cohort

1359-6446/06/$ - see front matter http://dx.doi.org/10.1016/j.drudis.2013.11.003

identification and analysis [14]. Although database specific, an EHR for an individual patient typically includes a unique patient id, clinical diagnoses (as standardized diagnostic codes), drug prescriptions, laboratory test results and, in some cases, lifestyle information, such as smoking, drinking and body mass index. Coverage of the underlying population is likely to be broad, and new linkages between databases are enhancing the ability to ascertain disease status. For example, the UK Clinical Practice Research Datalink (CPRD) primary care database contains anonymized health information for approximately 8% of the total UK population and can be linked to the UK Hospital Episode Statistics database [15].

Effective case identification algorithms can be developed that utilize EHR databases to identify cases of drug-induced liver injury. Work has been performed demonstrating the portability of such algorithms across different institutions [16] and work is ongoing to facilitate standardized implementation across databases in different countries (Ruigomez and Brauer, unpublished, for the IMI PROTECT group [17]). An example algorithm is provided in Fig. 1. Potential drug-induced liver injury cases are identified based upon specific diagnostic codes (routinely inputted by clinicians). The database is searched for liver test results within a specific time period

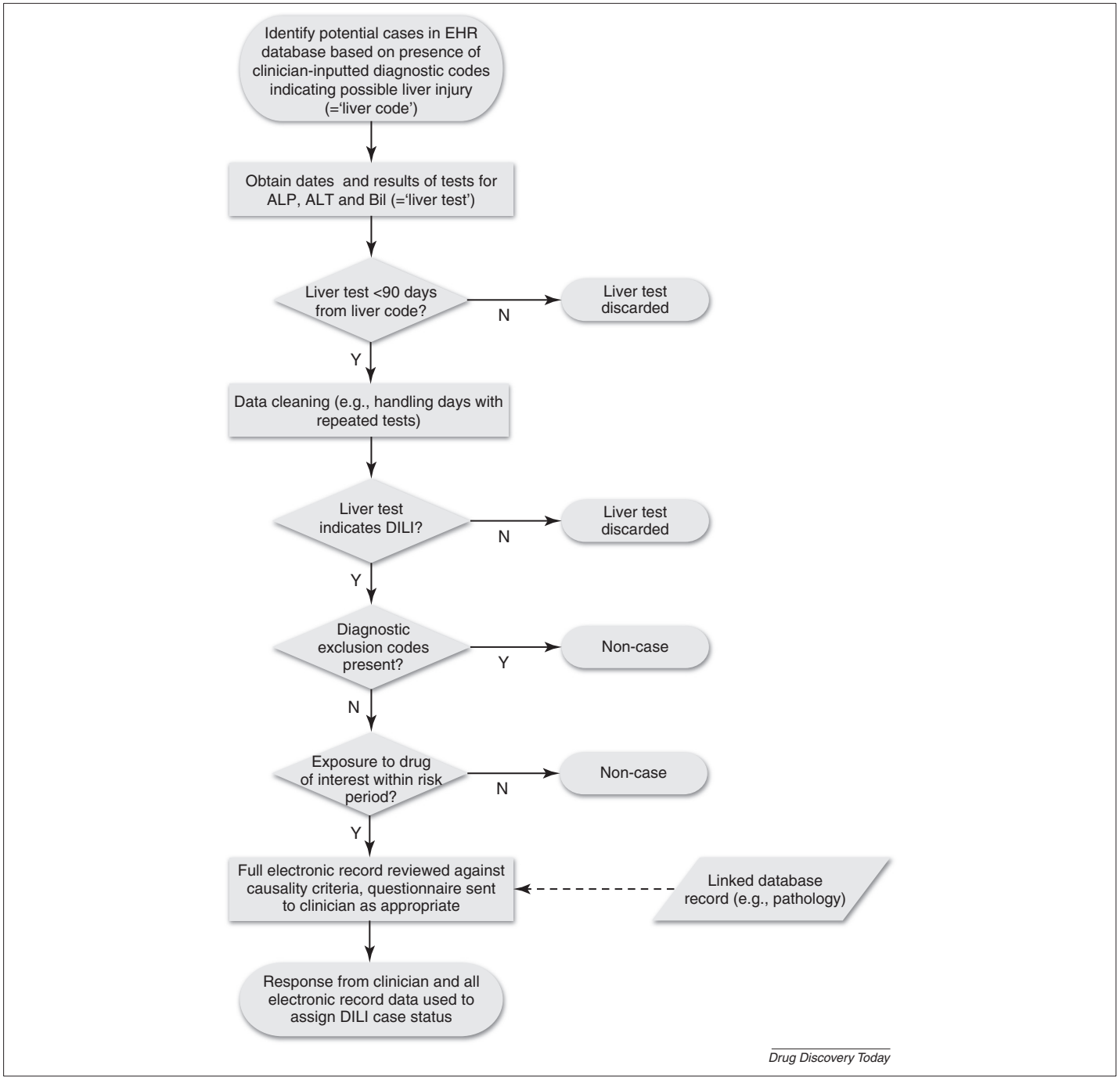

FIGURE 1

Example algorithm for identifying cases of drug-induced liver injury using a database of electronic health records. Drug-induced liver injuries (DILI) are injuries defined as ALT≥5 × ULN, ALP≥2 × ULN, or ALT≥3 × ULN with Bil>2 × ULN [13]. Causality criteria are as determined by international consensus [17].
Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; Bil, bilirubin; EHR, electronic health record.