SUPPLEMENTARY INFO ON SpotRM+

The overall goal of the novel expert database SpotRM+ is to deliver easy access to what is known about how certain chemical substructures are metabolically activated resulting in reactive metabolites (RM). Designed as a didactic tool, SpotRM+ with a solid knowledge base will hopefully lead to a deeper analysis of potential toxicity risks with planned test compounds and thus result in a more efficient lead selection and optimization. The basis of our selection is reviewed in a Perspective paper: A. Claesson & A. Minidis, *Chem. Res. Toxicol*. 2018.

The design of SpotRM+ is focused on delivering:

- A comprehensive compilation of structural features that are believed to be involved in the clinical adverse effects of certain drugs by formation of RMs.
- Unprecedented, facile access to the most relevant information on structural alerts and reported evidence of their involvement in causing side-effects.
- Continuous improvements in all aspects, such as supplementing data regarding new alerts, drugs, experimental compounds, and new references.

General considerations regarding drugs with toxicity issues

The number of drugs that have been on the market and then withdrawn due to *toxicity associated with their metabolism* is frustratingly large, well over one hundred. The dominant problem here is metabolic bioactivation to form *reactive metabolites* leading to alterations in proteins and DNA. Since this toxicity also leads to termination of candidate drugs in clinical trials early identification of compounds that do not exhibit a risky metabolic conversion cannot be sufficiently stressed (*cf*. reviews collated at website stoprm.org). It is our belief that SpotRM+ can contribute to facilitating the process of finding these drug structures.

The precise biochemical mechanisms of adverse events of many drugs, in particular no longer-marketed ones, have rarely been investigated. However, in many cases involvement of RMs might be considered when the clinical symptoms involve:

- hepatotoxicity
- effects on blood cells, e.g. agranulocytosis
- allergic reactions, including skin reactions and anaphylactic shock
- carcinogenicity
- drug-drug interaction (via mechanism-based inhibition, MBI, a.k.a. time-dependent inhibition, TDI)

Furthermore, many unknowns are involved in estimating the clinical consequences of RM formation since the reasons for discontinuation of clinical trials are often unknown to outsiders. Such trials also naturally produce limited amounts of safety data. Additionally, it is important to keep in mind that experts generally consider side-effects of licensed drugs to be underreported.

SpotRM+ tries to draw examples of drugs from the four first of the above-mentioned types of clinical outcomes. Naturally, our conclusions on drugs where the literature is scarce (as hinted, a very common situation) should be regarded as provisional.

Another growing set of compounds consists of those described in the medicinal chemistry/ pharmacology literature with *in vitro* data on adduct formation with trapping agents, foremost glutathione. Illustrative examples of such compounds are included where pertinent. Updates of

SpotRM+ will include even more examples of this type.

Classification of drugs is based on perceived severity of adverse drug effects

In the SpotRM+ database all alerts are linked to selected examples of the drugs that have the alert (or, rarely, a very similar substructure). For a rapid visual impression of the hazard associated with a given alert, the drugs are classified as *Red*, *Yellow*, *Green* or *Neutral*. Definitions of these classes are:

Red

The drug has shown clinical adverse effects that have a proven or highly probable association with bioactivation to RMs **or**, regarding preclinical compounds, experiments have shown extensive formation of RMs.
A drug used clinically has been withdrawn or bears regulatory warnings.
Examples: *sudoxicam, felbamate, amodiaquine, lamotrigine* or *ritonavir*.

Yellow

The drug contains an alert included in SpotRM+, and there have been some reports of adverse effects that have been discussed in terms of RM formation.
Preclinical compounds have displayed RM formation.
Examples: *mirtazapine, zolpidem, atorvastatin, tolmetin* or *phenazone*.

Green

Despite having triggered an RM alert, the drug has been used clinically without reported findings of adverse effects that can be associated with RM formation. The explanations for vindication include low daily dosage or a very low degree of metabolism that involves the relevant substructure. The last-mentioned case often depends on metabolism in other parts of the molecule.
Examples: *rivaroxaban, cefuroxime, aripiprazole* or *tolterodine*.

Neutral

In spite of having a RM alert in the structure the clinical information is insufficient to classify the drug into any of the other categories. This might be because the drug has only recently been introduced or is not widely used, and therefore sufficient safety data have not been accumulated.
Examples: *formoterol, fluorofelbamate* or *ranolazine*.

Bioactivation is required, chemical reactivity per se excluded

SpotRM+ focuses exclusively on drugs that are *bioactivated to form reactive intermediates*. There are quite a number of drugs on the market that give rise to toxic effects, including hepatotoxicity, by being *chemically reactive per se*. Examples range from beta-lactams to cytotoxic anticancer agents, for example busulfan. There is also current interest in making irreversible kinase inhibitors. Because of the difficult decisions in what to include, these are all currently excluded but may be considered for new versions or extra modules of SpotRM+.

Cytotoxic anticancer agents are largely left out

A large number of anticancer agents require metabolic bioactivation to achieve *cytotoxic therapeutic activity*, a principle that entails many well-known and feared side-effects. For the purpose of SpotRM+, which is to help spotting hidden alerts in "normal" drugs, this kind of overt toxicity is of minor interest and therefore the majority of these drugs have been excluded. Nevertheless, some anticancer agents may well offer guidance regarding what safety-focused medicinal chemists ought to keep out of their drug design tool-box. We have tried to incorporate this type of learning in SpotRM+.

Selection of alerts - focus on the less well-known

It is a delicate task to select the really useful substructures (RM alerts), which will be represented by a well-written SMARTS string for matching in target structures. The dilemma can be illustrated by including a string that corresponds to a benzene ring with nitrogen attached to it. This would encompass very many drugs. We have instead chosen to split this substructure class into many smaller classes; a few examples are indoles, o-methyl-benzeneamines, N-acyl-benzeneamines, benzimidazoles, and more. In this way we can fine-tune the examples that are most appropriate. The user can still search using as input benzene with attached nitrogen but as result may have to handle up to a hundred or more hits.

SpotRM+ is not encyclopaedic: the more widespread the knowledge of the hazard of a certain substructure is the less is the marginal gain to the user of the application of having it included. Many alerts in SpotRM+ are of this type, for example a nitroarene, which is more or less taboo in drug candidates of today though we choose to include several examples of this for the sake of reasonable comprehensiveness. Regarding the number of examples of drugs (in the red/yellow/green/neutral scale) we find less value in comprehensively listing in SpotRM+ all the drugs that are anilines (including "hidden" anilines), which have displayed hepatotoxicity. Besides, many of these are hugely obsolete. On the other hand, safe anilines and an explanation as to why they may be considered safe are certainly within the scope of SpotRM+.

To increase the usefulness of SpotRM+ we have made considerable efforts to include as alerts substructures that have not been mentioned explicitly as such in the literature but probably have contributed to adverse effects. In particular, our selection of alerts focuses on the less well-studied hazards of certain combinations of ring systems and substituents. Here you may find, for example a range of substructures that ultimately result in various types of *activated benzylic alcohols* such as derivatives of 2- or 4-methyl-benezeneamine, or five-membered heterocycles having methyl groups.

Since there is a huge lack of understanding of how adverse effects might be associated with RM formation, we have been generous with hints of potential compound liabilities and offer many hypotheses regarding less obvious reaction pathways based on sound chemical reasoning and guessing of routes of metabolism. Our sincere hope is that our users will be involved in the continuous improvement of SpotRM+ by communicating their points of view to us.

Delimiting an alert/toxicophore is a delicate task as well. We hope to work together with the users of SpotRM+ in a process to refine the SMARTS definitions and the linking of alerts to the most illustrative example drugs.

A comment on irreversible CYP enzyme inhibition

Many drugs interfere with the metabolism of other drugs, causing drug-drug interactions. This can take place in several ways, but the most common mechanism, which is also relevant in the current context, is irreversible inhibition of a CYP enzyme that is responsible for metabolising another drug. The principle mechanism of such inhibition is via a reactive species that is a RM formed *in situ* at the enzyme's active site. Most drugs that have a reported significant evidence of this mechanism are included in SpotRM+, since effects collateral with RMs formed *in situ* may also cause other types of toxicity. However, there is a need to systematically include even more drugs that inhibit CYPs; future updates of SpotRM+ will gradually fill the gap in order to ensure comprehensiveness.

Other aspects

The user of this tool should duly appreciate the vast limitations in current knowledge of the ways most drugs are metabolized, which often proceed via multiple sequential and branching reactions. Hence there is often great uncertainty regarding the nature of the RMs and their respective roles in causing adverse reactions by reacting with proteins and DNA. Adding to the limited possibilities of describing cause and consequences in this field is the individual susceptibility to RMs that lies behind idiosyncratic drug reactions.