Minireview

Recommendations on the Use of Multiple Labels in Human Mass Balance Studies

Filip Cuyckens, Mette G. Hvenegaard, Kenneth C. Cassidy, Douglas K. Spracklin, Alexander D. James, Mette L. Pedersen, Graeme Scarfe, David S. Wagner, Katrin Georgi, Simone I. Schulz, Hanno Schieferstein, Inga Bjornsdottir, Andrea A. Romeo, Georges Da Violante, Stefan Blech, Patricia Moliner, and Graeme C. Young

Janssen R&D, Beersel, Belgium (F.C.); H. Lundbeck A/S, Copenhagen, Denmark (M.G.H.); Eli Lilly and Company, Indianapolis, Indiana (K.C.C.); Pfizer Inc., Groton, Connecticut (D.K.S.); Novartis, Basel, Switzerland (A.D.J.); Biopharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden (M.L.P.); Sosei Heptares, Cambridge, United Kingdom (G.S.); AbbVie, North Chicago, Illinois (D.S.W.); Bayer AG, Wuppertal, Germany (K.G., S.I.S.); The Healthcare Business of Merck KGaA, Darmstadt, Germany (H.S.); Novo Nordisk, Maaloev, Denmark (I.B.); Roche Pharma Research and Early Development, Basel, Switzerland (A.A.R.); Servier, Gif-sur-Yvette, France (G.Da.V.); Boehringer-Ingelheim Pharma GmbH & Co. KG, Biberach, Germany (S.B.); Sanofi, Montpellier, France (P.M.); and GSK Research & Development Ltd., Stevenage (G.C.Y.)

Received September 18, 2023; accepted November 27, 2023

ABSTRACT

The administration of radiolabeled drug candidates is considered the gold standard in absorption, distribution, metabolism, and excretion studies for small-molecule drugs since it allows facile and accurate quantification of parent drug, metabolites, and total drug-related material independent of the compound structure. The choice of the position of the radiolabel, typically 14C or 3H, is critical to obtain relevant information. Sometimes, a biotransformation reaction may lead to cleavage of a part of the molecule. As a result, only the radiolabeled portion can be followed, and information on the fate of the nonlabeled metabolite may be lost. Synthesis and administration of two or more radiolabeled versions of the parent drug as a mixture or in separate studies may resolve this issue but comes with additional challenges. In this paper, we address the questions that may be considered to help make the right choice whether to use a single or multiple radiolabel approach and discuss the pros and cons of different multiple-labeling strategies that can be taken as well as alternative methods that allow the nonlabeled part of the molecule to be followed.

SIGNIFICANCE STATEMENT

Radiolabeled studies are the gold standard in drug metabolism research, but molecules can undergo cleavage with loss of the label. This often results in discussions around potential use of multiple labels, which seem to be occurring with increased frequency since an increasing proportion of the small-molecule drugs are tending towards larger molecular weights. This review provides insight and decision criteria in considering a multiple-label approach as well as pros and cons of different strategies that can be followed.

Introduction

Incorporation of a radiolabel into the small-molecule asset under development is seen as the gold standard approach in establishing quantitative information on the absorption, distribution, metabolism, and excretion (ADME) of the drug (Penner et al., 2009; Spracklin et al., 2020). Also, for health authorities, the radiolabel human ADME study holds a central place in the understanding of the disposition and elimination of a drug in development. The European Medicines Agency has provided guidelines (https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-drug-interactions-revision-1_en.pdf), and the US Food and Drug Administration recently issued draft guidelines (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-pharmacology-considerations-human-radiolabeled-mass-balance-studies) around the conduct of this study. They state that a radioactive label in a chemically and metabolically stable position should be used and that “in some cases two separate labeling positions have to be used to follow the fate of the drug under investigation.”

Radiolabeled carbon (14C) is the most frequently used radioisotope label in human ADME studies. As an alternative to 14C, tritium (3H) labeling can be used as it is often more easily introduced. It provides higher sensitivity for radiodetection due to its higher specificity. However, it is potent and/or a peptide or larger-molecular-weight drug. Therefore, alternative methods that allow the nonlabeled part of the molecule to be followed.
due to its higher risk of biologic instability, most often arising from oxidative cleavage with loss of the label via tritiated water, and potential nonuniform distribution of \(^3\)H label(s), it is more often used in earlier stages, i.e., in discovery, whereas \(^14\)C labels are usually preferred in later stages of drug development (Isin et al., 2012; Krauser, 2013).

In any case, the position of the radiolabel needs to be carefully considered and ultimately selected to obtain the intended information from the ADME study. Also, for \(^14\)C-labeled compounds, there are instances, sometimes foreseen and sometimes unexpected, where the single-label position approach will not provide the required information as the label is lost through bond cleavage via degradation or metabolism. Thus, the routes and rates of the ADME processes under investigation, including that of significant constituent parts of the drug, are potentially not traced either qualitatively or quantitatively, triggering the use of more than one radiolabel in an ADME study. Drug metabolism scientists are having these discussions around multiple-labeling approaches with increasing frequency since an increasing proportion of the small-molecule drugs are tending toward being of larger molecular weight. This is in part due to the advent of new technologies, making these molecules druggable, and new modes of action such as protein degraders, also known as proteolysis targeting chimeras (PROTACs) (Goracci et al., 2020; Pike et al., 2020). Also, for large molecules, such as peptides and oligonucleotides, the selection of an optimal labeling position or use of multiple labels can be an important but difficult discussion topic. Although some of the aspects discussed in this paper might be relevant, we consider large molecules out of scope since they require a different and case-by-case approach (endogenous versus exogenous building blocks, potential recycling of the substructure bearing the label, etc.). Often, a radiolabeled or nonlabeled ADME study will not provide the right answers and, thus, is not a relevant study to perform for these modalities.

In this paper, we address the questions that should be considered to help make the choice between the use of a single radiolabel or multiple-labeling approach for the animal (where appropriate) and ultimately the human ADME study. Examples of strategies and approaches that have been taken to address this issue will be discussed as well as the pros and cons of these different approaches. An alternative to multiple labeling the drug is the use of an analytical method that can follow the nonlabeled part of the drug, and this will also be briefly explained. The recommendations provided are a result of discussions within the European Federation of Pharmaceutical Industries and Associations drug metabolism and pharmacokinetics Network–sponsored human ADME paradigm shift(s) working group.

When More Than One Label Is Recommended

As mentioned above, there are instances where the single-label position approach will not provide the required information as the label is lost through cleavage and significant constituent parts of the drug are potentially not traced, triggering the use of more than one radiolabel in a human ADME study. The number of labels needed is evaluated on a case-by-case assessment, which is based on whether cleavage of the molecule into significant parts is expected. This evaluation will be based on prior experience, including knowledge within drug metabolism and experimental building blocks, potential recycling of the substructure bearing the label, etc.). Often, a radiolabeled or nonlabeled ADME study will not provide the right answers and, thus, is not a relevant study to perform for these modalities.

![Fig. 1. Amide hydrolysis pathways observed for (A) pomotrelvir [adapted from (Yang et al., 2023)] and (B) DS-8500a [adapted from (Makino et al., 2022)]. *\(^14\)C labeling position.](image-url)
in vitro and preclinical in vivo data and/or data obtained from the cold first-in-human study. Yang et al. (2023) opted for a dual radio-labeling approach for the human mass balance of pomotrelvir since they anticipated significant hydrolysis of one of the amide bonds based on metabolite profiling in preclinical species and in the first-in-human study (Fig. 1A). Both cleavage metabolites were shown to be major circulating metabolites, exceeding 10% total drug-related material. The dual labeling also facilitated observation of some important early-eluting secondary metabolites of this cleavage pathway that were not picked up in the cold first-in-human study due to their poor ionization efficiency in mass spectrometry (Yang et al., 2023). Makino et al. (2022) described how dual radio-labeling was triggered for the human ADME study of DS-8500a based on the observation of oxazolidine ring-opened and ring-cleaved metabolites (Fig. 1B) as a major pathway in rats in combination with the presence of the same metabolites in human in vitro systems. In the human mass balance study, another hydrolysis pathway showed to be more pronounced, resulting in the loss of an unlabeled 2-amino-propan-1-ol moiety (Fig. 1B). Even if this information had been available or anticipated prior to the study, the choice for the position of the second label would probably have persisted because it was placed at the core of the molecule and covers the oxazolidine ring cleavage metabolites, whereas 2-amino-propan-1-ol is a small polar aliphatic moiety bearing no real structural alert. Adding a third label, on the other hand, would have greatly overcomplicated the study. Absence of cleavage of the drug in animal species does not exclude that the cleavage will occur in humans. To this end, Stypinski et al. (2020) observed a human-specific cleavage in their metabolite profiling study of 14C-Iloratinib, which resulted in the need for a repeat human ADME study with the radiolabel in a different position to the original study.

If it is concluded, based on the considerations mentioned above, that the molecule can be cleaved into two or even three significant parts, the need for more than one label should be evaluated. If the cleavage is expected to be a significant pathway in humans as eventually determined via metabolite profiling in excreta, more than one label is recommended. The authors of this paper suggest using an expected threshold of >25% of metabolic clearance as a compromise between what is practically feasible and what is acceptable from a risk perspective. However, even if the extent of the cleavage pathway is low in excreta or maybe not even observed, this does not exclude that a cleavage product may be observed in plasma and potentially be Metabolites in Safety Testing (MIST) relevant (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/m3r2-nonclinical-safety-studies-conduct-human-clinical-trials-and-marketing-authorization; Leclercq et al., 2009; Luffer-Atlas et al., 2023). It has previously been the case that metabolites not observed or only observed to a low extent in excreta were MIST relevant, i.e., exceeding 10% of total drug-related material in plasma (Glaenzel et al., 2018). In contrast to excreta, a “mass balance” approach cannot be made in plasma as the cleavage products would likely have different volumes of distribution and thus be present to different extents in plasma. Furthermore, cleavage products are often more hydrophobic than the oxidized parent drug or unchanged parent itself. If this is the case, they may circulate to a higher extent, increasing the likelihood of becoming MIST relevant. Therefore, the low abundance or apparent absence of the cleavage products in excreta cannot alone be used to rule out that more than one label is needed. Since the assessment of metabolites exceeding 10% of total drug-related material in plasma is a widespread approach in first-in-human studies, this information is often available prior to the planning of the human mass balance study. It is, however, critical that suitable analytical approaches are applied since the mass spectrometry response of cleavage products can be very different from that of the unchanged drug, which can lead to an underprediction of their abundance (Leclercq et al., 2009; Luffer-Atlas et al., 2023).

If the metabolism pathway of the cleavage products is unknown (not already described in literature), more than one label is recommended. However, if one of the cleavage products is not expected to be substantially metabolized further, this moiety could potentially be traced by liquid chromatography–mass spectrometry quantification and, with that, avoid the need for labeling of that part of the parent molecule (please refer to Analytical Techniques Enabling Detection of Nonlabeled Biotransformation Products). Furthermore, if one of the cleavage products is a known endogenous innocuous molecule and if the expected concentrations and amounts formed are not significantly higher than the endogenous level, labeling of this part is not considered necessary. Additionally, if one of the cleavage products generates a metabolite with previously documented metabolism, multiple labeling should not be considered. However, if the cleavage product is a known molecule with a potential safety risk, a decision on the labeling or targeted nonlabeled approaches should be considered case by case based on 1) knowledge of the clearance and toxicity profile of this entity, 2) anticipated exposure levels, 3) safety coverage in nonclinical species, 4) proposed therapeutic indication, etc. The loss of formaldehyde, resulting from N- or O-demethylation reactions, is a typical example of where the release of a toxic agent does not require further follow-up since formaldehyde is omnipresent, and the additional exposure due to demethylation of drug will be negligible. This remains true as long as the dose used is not excessive and the attendant formaldehyde exposure remains low against the context of background exposures.

Another consideration when deciding on the number of labels to include is the feasibility with respect to synthesis as well as stability of the radiolabeled compounds. If, for instance, the second label is unstable due to radiolysis (i.e., chemical decomposition due to ionizing radiation), or the second 14C-label is only possible on a carbon that is expected to decarbonylate and thus be released as 14CO2, performing the studies with one label and potentially combining data provided by nonlabeled approaches may be the only option. A summary of the main criteria to consider for a multiple labeling approach are highlighted in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Criteria to consider a multiple-labeling approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Cleavage of the molecule into significant parts is expected.</td>
</tr>
<tr>
<td>- Cleavage pathway is expected to exceed 25% of metabolic clearance and/or exceed 10% total drug-related material in plasma.</td>
</tr>
<tr>
<td>- The metabolism pathway of the cleavage product(s) is unknown, or the metabolite is a known endogenous molecule and the expected concentrations have the potential to exceed known safety margins.</td>
</tr>
<tr>
<td>- The nonlabeled part(s) cannot be followed by nonlabeled analytical techniques.</td>
</tr>
<tr>
<td>- Use of multiple labels is technically feasible.</td>
</tr>
</tbody>
</table>

### Analytical Techniques Enabling Detection of Nonlabeled Biotransformation Products

Liquid chromatography–mass spectrometry (LC-MS/MS)-based bioanalytical assays can be applied to quantify metabolites formed by metabolic cleavage with loss of the radiolabel. A recent example has been described by Piel et al. (2023), where amide cleavage was found to be one of the major metabolic pathways in the human
metabolism of the oral factor Xa inhibitor asundexian. The resulting nonlabeled metabolites, including the only major human metabolite (>10% of total drug-related material in plasma), were quantified using LC-MS/MS.

This analyte-specific assay approach may, however, be useful when only one or a few nonlabeled metabolites need to be monitored since full structure identification and synthesis of authentic standards is required. Bioanalytical methods need to be developed in multiple matrices, and all of this requires significant amounts of resources. Furthermore, this specific analyte assay approach does not facilitate tracing of downstream metabolites (i.e., secondary or tertiary) to the nonlabeled moiety. However, powerful qualitative tools exist, such as high-resolution mass spectrometry used in combination with untargeted software-assisted metabolite detection to extract metabolite peaks from the endogenous background. In vitro and/or in vivo nonclinical data can be used to identify the main downstream metabolites of the nonlabeled cleavage product and assess the need for their synthesis in preparation for the human ADME study. It should, however, be noted that there is a risk that some metabolites with low response by mass spectrometry (e.g., very polar small compounds) or metabolites being present at low concentrations (e.g., low-dose drugs) may not be discovered as there is no radioactive tracer present to guide where and what to look for. Furthermore, since the radioactivity offers a structure-independent detector response, the ability to quantify the metabolites is lost for the moieties not containing the radiolabeled part and with that the understanding of the importance of each metabolite in terms of concentrations or amounts.

So-called “universal detectors,” such as evaporative light scattering detection and charged aerosol detection, lack the necessary selectivity and sensitivity to quantify drug metabolites in complex endogenous matrices.

Since many drug candidates contain one or more halogens, i.e., close to 30% of recently approved drugs (Benedetto Tiz et al., 2022), fluorine-19 nuclear magnetic resonance spectroscopy (19F-NMR) and inductively coupled plasma–mass spectrometry (ICP-MS) can be considered to follow cold metabolites where the label is lost (Leclercq et al., 2009). A good correlation with radioactive quantification was demonstrated for both 19F-NMR (Lenz et al., 2002a,b; James et al., 2017) and ICP-MS (Jensen et al., 2005; Cuyckens et al., 2008; Izmer et al., 2012). Chlorine, bromine and iodine can all be analyzed by ICP-MS, and this is probably the preferred technique for metal-containing drugs (Gammelgaard and Jensen, 2007). More recently, its potential for the analysis of fluorine-containing compounds has also been diligently explored (Zhu et al., 2017; Redeker et al., 2022). Although ICP-MS, unlike 19F-NMR, benefits from relatively easy coupling with liquid chromatography, its application to fluorine measurements remains limited due to its poor sensitivity for this element.

Since both techniques provide a structure-independent response, not only can nonlabeled cleavage products and their downstream metabolites be quantified but such techniques can also be considered worthy alternatives to the use of a radiolabel as illustrated by their application in human mass balance studies (Meermann et al., 2012; Pearson et al., 2019; Singh et al., 2022). In addition to avoiding the disadvantages of working with radioactivity (radiosynthesis, exposure to radioactivity, radioactive waste), these approaches allow human metabolism data to be obtained in ongoing clinical studies (e.g., first in human). This also facilitates easier generation of steady-state data, not only for the parent drug but also for metabolites, which is of particular importance in the case of time-dependent pharmacokinetics. The relatively limited sensitivity of 19F-NMR and potential interference of environmental halogen-containing compounds are still the main bottlenecks for more routine application of both techniques in drug metabolism studies. The potential interference of environmental halogen content also makes it almost impossible to directly measure untreated samples, such as feces or feces homogenates, to allow assessment of extraction efficiency. Furthermore, it should be taken into account that, similar to a radiolabel, the halogen(s) can also be lost through metabolism as was addressed in the nirmatelvir study with the measurement of a 19F-NMR-silent metabolite by LC-MS/MS (Singh et al., 2022).

How Can a Study with Administration of Multiple Labels Best Be Performed?

If there is no other option but to use multiple radiolabels in the human mass balance study, different approaches can be applied: dosing 3H- and/or 14C-labeled compounds separately or as a mixture or even combining labels in one molecule. Table 2 provides an overview of the different strategies that can be followed with their pros and cons.

The most straightforward, and by far the technically easiest, approach is the separate dosing of the different labels to different subjects. The potential downside of this method is that more subjects may need to be included, although the difference might be marginal, e.g., 4 + 4 subjects dosed when two labels are needed compared with six or seven (to warrant six evaluable subjects) in a study applying one label. Additional preclinical work to support separate dosing will be required, including, for each label, a rat mass balance study and a rat quantitative whole-body autoradiography study, which is necessary for dosimetry purposes unless a 14C-microtracer strategy is used in the human study (Young et al., 2023). When a study dosing a 1:1 mix of 14C-labeled compounds, or a single molecule with two 14C labels, is chosen, then one must rely on metabolite identification to assign the specific activity for each metabolite. For example, if a molecule is a 1:1 mixture of two 14C labels or a single parent molecule with two 14C labels with a specific activity of 2 MBq/mmol and is cleaved, each metabolite resulting from the cleavage would only have a specific activity of 1 MBq/mmol. If the weight equivalents of the metabolites are based on the specific activity of parent (2 MBq/mmol), then the actual concentration of a cleaved metabolite is 2 times what is measured. This is true if both 14C labels have the same specific activity, which is usually not the case, thereby further complicating the calculations. For the same reason, the use of multiple labels in one molecule is highly discouraged since the assessment of the individual specific activity of the different labels in one molecule is extremely challenging. The stability of the radiolabels might also be different. In addition, if there are many unidentified minor metabolites in circulation without assignment as intact or cleaved metabolites, the calculations of percent of circulating drug-related material may be distorted. Similarly, unidentified metabolites in excreta may hamper a correct assessment of the assignment of percent clearance pathways. Furthermore, conducting a study with a true mixed 14C and 3H dual label in the same molecule is even more complex. Energy windows of the radiodetector can be adjusted to mitigate the overlap and contribution of the 3H channel in the 14C measurement with a small negative impact on 14C sensitivity. The impact of the 14C channel on the 3H measurements can, however, not be canceled because radioactive measurement from the higher energy 14C spills (crosstalk) into the lower energy 3H channel, resulting in an overestimate of 3H counts. Separate liquid chromatography injections of the single-labeled standards are required to define the specific activity of each label and the degree of crosstalk of the 14C label in the 3H measurement (Shaffer and Langer, 2007). This approach was applied in the human ADME study of ziprasidone, where for the simultaneous monitoring of codosed 3H- and 14C-labeled compounds, efficiencies of 37% for 3H and 55% for 14C were used, with a compensation for 14C spillover into the 3H window of 31% (Prakash et al., 1997).
Separate specific activity measurements can, of course, not be performed in the case where both \( ^3\text{H} \) and \( ^1\text{C} \) labels are contained in the same molecule. Therefore, this approach is highly discouraged and likely also one of the reasons we could not find any human ADME studies using a mix of \( ^3\text{H} \) and \( ^1\text{C} \) labels in one molecule. The only scientifically sound way to selectively analyze both \( ^3\text{H} \) and \( ^1\text{C} \) levels in this case is by applying accelerator mass spectrometry (AMS) since it can selectively analyze \( ^3\text{H} \) and \( ^1\text{C} \) concentrations (Roberts et al., 1994), but a lower sensitivity of AMS for \( ^3\text{H} \) detection should be taken into account. However, AMS cannot be applied for quantitative whole-body autoradiography, and therefore the radiation burden calculation will remain very complicated (if possible at all) due to different nuclides and radiation intensities.

**Conclusions**

When metabolic cleavages are possible, the incorporation of more than one radiolabel helps to ensure that all main metabolic clearance pathways and potential major circulating metabolites are identified. However, due to the complexity involved, before committing to a multiple radiolabel study it is best to consider all other options, such as alternative positioning of the radiolabel, following the nonlabeled part of the molecule via other suitable analytical techniques and whether there really is a need to follow a small cleavage product.

If there are no other options but to use more than one radiolabel in the human mass balance study, a study design involving dosing of the different labels separately to different participant cohorts is preferred. Although this strategy may increase costs and the number of subjects exposed to radiation, it will significantly reduce analytical complexity and ensure accurate radioactive quantification of fractions excreted and of circulating metabolites for individual metabolites. Additionally, there may be efficiencies to be gained through adopting a multiple-labeling approach in the first instance where appropriate versus having to conduct a repeat human ADME study later in development.

**Data Availability**

This article contains no datasets generated during the current study.
References


Cassidy, Spracklin, James, Pedersen, Scarfe, Wagner, Georgi, Schulz, Schieferstein, Bjornsdottir, Romeo, Da Violante, Blech, Moliner, Young.

Cuyckens, Hvenegaard, Cassidy, Spracklin, James, Pedersen, Scarfe, Wagner, Georgi, Schulz, Schieferstein, Bjornsdottir, Romeo, Da Violante, Blech, Moliner, Young.

Leclercq L, Cuyckens F, Mannens GS, de Vries R, Timmerman P, and Evans DC (2009) Which New chemical entities have been approved in the last ten years and are now on the market? Curr Opin Drug Discov Devel 12:616–626.


