Methodological approaches for personalised medicine: a series of scoping reviews

Protocol V.2– 29 April 2020

Authors

PERMIT WP 2 group

Methods team

Rita Banzi, Chiara Gerardi, Maddalena Fratelli
Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy

Review teams:

Teresa Torres, Josep Maria Haro Abad, Albert Sanchez Niubo
Parc Sanitari San Joan de Déu, San Boi del Llobregat, Barcelona, Spain
Enrico Glaab
Luxembourg Centre for Systems Biomedicine, Université du Luxembourg, Luxembourg
Emanuela Oldoni, Florence Bietrix, Vibeke Fosse, Emmet McCormack
European infrastructure for Translational Medicine(EATRIS), Amsterdam, The Netherlands and University of Bergen, Bergen, Norway
Raphael Porcher, Cecilia Superchi
Centre de Recherche Épidémiologie et Statistiques Université de Paris, Université Paris Descartes, Paris, France

PERMIT coordination

Paula Garcia, Jacques Demotes-Mainard
European Clinical Research Infrastructure Network (ECRIN), Paris, France

Protocol reviewed by WP2 contributors:

Lorena San Miguel, Frank Hulstaert
Belgium Health Care Knowledge Centre, Bruxelles, Belgium
Alexander Grundmann, Rainer Girgenrath
Deutsche Zentrum für Luft- und Raumfahrt, Köln, Germany
Luis M. Sánchez-Gómez, Mar Polo-de Santos, Iñaki Imaz-Iglesia, Setefilla Luengo
Instituto de Salud Carlos III, Madrid, Spain

Table of Contents

Introduction

The concept of personalised medicine is going to impact how pharmacological treatments are discovered and developed, how patients are diagnosed and treated, and how health care systems allocate their resources to maximize patient benefits.
Personalised medicine may be considered an extension of traditional approaches to understanding and treating disease. Ideally, it could serve to take clinical decisions based on a patient's profile (often molecular, but the concept is broader) to minimise harmful side effects, ensure a more successful outcome, and possibly help contain costs compared with a "trial-and-error" approach to disease treatment (1).

Personalised medicine stems on the broad concept that managing a patient's health should be based on the individual patient's specific characteristics, including age, gender, height/weight, diet, environment, etc. Different understandings of personalised medicine exist, in which three main positions can be identified (2):
(a) personalised medicine is not a new concept as medicine has always been individualized;
(b) personalised medicine is holistic health care, centred around the needs of the individual patient;
(c) personalised medicine is treatment targeted at stratified subgroups (e.g. pharmacogenetics).

Even when the focus is restricted to the third position, there is not a unique definition of personalised medicine, nor a straightforward terminology to define this concept. While "personalised" emphasizes the notion of individualized— "this is exclusively designed for you", other more scientifically rigorous terms such as *stratified medicine* refer to the identification of groups or strata of patients with specific molecular characteristics or other determining factors which predict susceptibility to disease, disease prognosis, and/or response to therapy. Some authors suggested that rather than considering personalised medicine as a precise scientific concept, it should be understood as an open and negotiable ideal that accounts for a plurality of visions, depending on people, reasons and interests behind these alternative conceptions (3).

Regarding the terminology, in the European context, the term **personalised medicine** is preferred, as this term best reflects the ultimate goal of effectively tailoring treatment based on an individual's 'personal profile', as determined by the individual's genetic and phenotypical characteristics. Other terms are widespread, for instance stratified medicine, mainly used in the UK, or precision medicine mostly used in US and broadly referred to the 4 P (preventive, predictive, personalised and participatory) medicine. While there may be small nuances in the literal meanings of these terms, they usually refer to the same concept when applied in practice (4).

A recent review reported that the literature about personalised medicine usually refers to two different semantic approaches. Firstly, patients' stratification, that is grouping individual patients in subpopulation according to their probability to have a therapeutic benefit from a drug or regimen. Secondly, treatment tailoring, that is the individual status of a patient (i.e., disease characteristics or subject's genotype/phenotype) is the rationale basis for drug choice (5).

Box 1 reports a collection of definitions, along with their references.

A broad community of stakeholders, including funders and professionals involved in medical research and care, are increasingly concerned with ensuring that the right patient receives the right therapy, at the right dose and at the right time. The identification of markers of mechanistic pathways or multiple variables characterising clusters of subjects that might inform meaningful disease stratification may have different clinical applications in the context of personalised medicine. Broadly, stratification may be applied at the diagnosis level (e.g., to identify a particular pathophysiological/clinical stratum within a heterogeneous patient population for diagnostic purposes), to predict disease course (prognostic value), the development of a disease (predictive value), or the response to therapy (theragnostic value).

Regardless of the application, any approach to personalised medicine should undergo different phases: discovery, validation and definition of usefulness from a clinical perspective. Robust methodological approaches are needed to deal with the complexity and heterogeneity of the process, as well as the range of possible applications to stratification using multidimensional data (what is meant by “molecular profiling” among other terms).

Personalised medicine research

This series of scoping reviews will map the general concept of methods for personalised medicine, to set the basis for the discussion on robustness and reproducibility of personalised medicine development programmes. The final goal is the identification of standards and needs in terms of methodology of data generation, management, analysis and interpretation to improve clinical studies in personalised medicine.

The group of authors agreed on a common operational definition of **personalised medicine research**: a set of comprehensive methods, (methodological, statistical, validation or technologies) to be applied in the different phases of the development of a personalised approach to treatment, diagnosis, prognosis, or risk prediction. Ideally, robust and reproducible methods should cover all the steps between the generation of the hypothesis (e.g., a given stratum of patients could better respond to a treatment), its validation and pre-clinical development, and up to the definition of its value in a clinical setting.

The process leading from the hypothesis to the clinic is complex and not always linear. The Medical Research Council in UK recently developed a framework for the development, design and analysis of stratified medicine (6) that is structured in six themes:

Theme 1: Framing the Question/Defining the Population
Theme 2: Designing Stratum Discovery Studies; selecting variables, defining response and powering
Theme 3: Assay Design; managing complexity and variability
Theme 4: Defining Strata; data integration, linkage to existing knowledge, linkage to outcome
Theme 5: Stratum Verification
Theme 6: Progression Towards Clinical Utility

Any attempt for classifying the phases of personalised medicine may appear as an oversimplification. However, a typical research programme in personalised medicine would include: first a stratification cohort (in many cases a retrospective study reusing data and biosamples from existing cohorts) with extensive multimodal data on which stratification algorithms are run, then a validation cohort, normally prospective, that assesses the reproducibility, robustness and validity of the clustering in another sufficiently large patient sample. Thirdly, a translational step is often necessary. In some cases, the use of pre-clinical models (cellular, in-silico, organoid) might be useful to give confidence in the allocation of patients to specific treatment arms as identified through clustering. Alternatively, the multi-omics profiles from clinical samples can lead to the identification of new disease categories, prediction of disease prognosis, exploration of drug sensitivity and dose selection. Finally, treatment options should be tested in the subgroups of patients in the context of clinical studies, ideally randomised clinical trials, to generate evidence informing regulatory, clinical and coverage decisions.

However, many alternative pathways can be proposed. In some case, the stratification provides detailed information on the mechanism of disease and strong indications on the treatments to be tested in each patient cluster. This is for instance the case where identification of driver somatic mutations in cancer cells suggests the targeted treatment to be tested. In other cases, the stratification cohort includes data on response to an established treatment, making the translational step less necessary. Research programmes may be limited to the stratification step, in particular when no treatment is available – this is the case for instance for taxonomy studies in

neurodegenerative disorders, aiming at identifying homogeneous clusters of patients. In any case, personalised medicine research is a complex programme, with multiple steps and lasting many years.

We consider out of the scope of this review the methods used for the clinical implementation of personalised medicine, the manufacturing and use of individualized treatments, and the pragmatic approach to individual patient care, such as n-of-1 trials.

Considering this framework outlined by Figure 1, the scoping reviews will approach **personalised medicine research** focusing on four main phases:

- Methods for stratification and validation cohorts
- Methods for machine learning applied to stratification
- Pre-clinical methods for translational development of stratified therapies and treatments selection
- Methods for clinical trials in personalised medicine

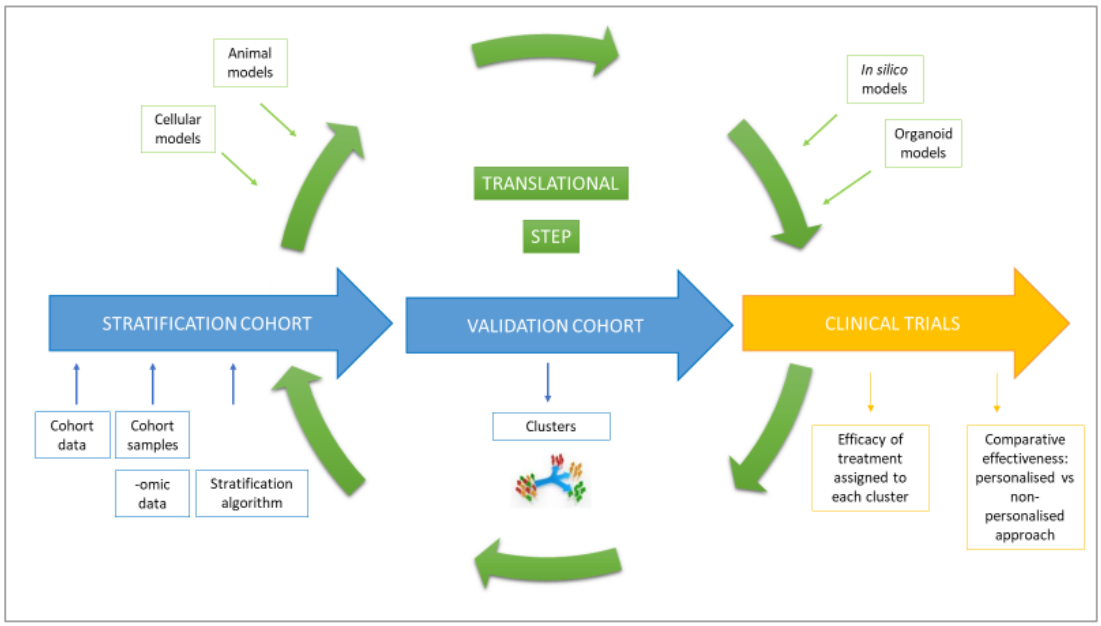

Figure 1: Main steps in personalised medicine research programmes

Review methods

We aim to perform a set of scoping reviews investigating various aspects of the methodology applied in personalised medicine research programmes as outlined in the Scope & Research Questions section.

Scoping reviews are used to present a broad overview of the evidence pertaining to a topic; they are useful to examine areas that are emerging, to clarify key concepts and identify gaps. Scoping reviews have great utility for synthesizing research evidence and are often used to map existing literature in a given field in terms of its nature, features, and volume. They differ from standard systematic reviews that are usually aimed to answer a specific question or series of questions according to a rigid set of a priori eligibility criteria. Scoping reviews have a broader approach, generally with the aim of mapping literature and addressing broader research questions. Due to the iterative nature of scoping reviews, deviations from the protocol are expected, differently from what happens in systematic reviews. Anyway, the discrepancies from the protocol will be clearly detailed and justified in the 'Methods' section of the scoping review report, if and when they occur.

To ensure the transparency and reproducibility of the review process, we will follow the methodological guidance for the conduct of scoping reviews suggested by the Joanna Briggs Institute (7, 8). The main steps of the process are summarised in Figure 2.

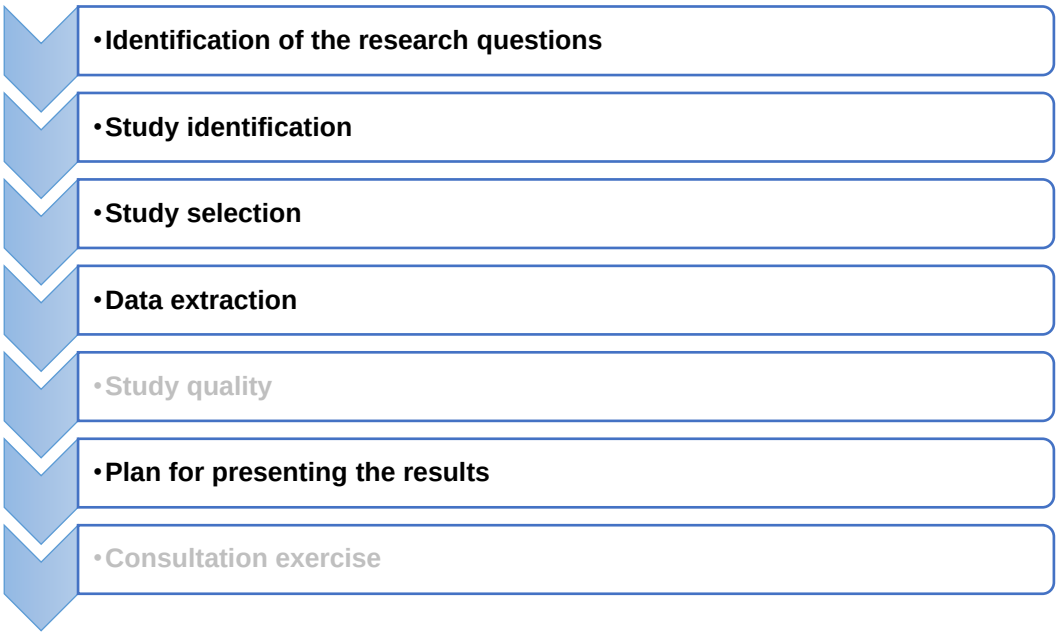

Figure 2: Main steps in the preparation of scoping review (in grey optional steps).

This overall process will be applied to the four themes outlined below by a dedicated review team supported by a methods team. Each step may require small adaptations given the nature of the research questions and scope defined and the type of literature/data that will be retrieved.

The four reviews will be part of a unique report covering the different aspects of methodology to inform the gap analysis and the subsequent phases of the PERMIT project (https://permit-eu.org/).

Identification of the research questions

The first step of any scoping review is to define the objective and research questions of interest. For the purpose of these reviews, four main aspects of the general concept of personalised medicine have been identified and will be the focus of this analysis:

- Methods for stratification and validation cohorts
- Methods for machine learning applied to stratification
- Pre-clinical methods for translational development of treatment options and treatments selection
- Methods for clinical trials in personalised medicine

Through several rounds of joint discussion and one face-to-face meeting (Paris, Jan 24, 2020), the four review teams had clarified the scope and defined research questions to the purpose of the scoping reviews. As the four topics are connected, this step also served to harmonise the four parts and avoid possible overlaps.

The outcome of this exercise is reported in the following section Scope & Research Questions.

Scope & Research Questions

1. Methods for stratification and validation cohorts

The scoping review will focus on:

- The characteristics of cohorts that have been used for patient stratification or validation of patient clustering obtained through stratification cohorts. Stratification cohorts of patients are used to create the clustering, and validation cohorts of patients are used to assess the reliability (robustness, reproducibility, etc.) of patient clustering.
- The different methods and tools used in design and management of stratification and validation cohorts (especially complex in multimodal approaches) to understand their limitations.

The review will not be restricted to a given type of data for stratification, i.e., genetic, metabolomics, gene expression, genomic, neuroimaging, etc.

General papers that describe methods and tools in the design and management of stratification and validation cohorts will be assessed irrespective of the diseases field. Case examples of biomarkers or multimodal data profiling in different medical fields and coming from different sources (omics, neuroimaging, genetics…) will be also analysed to explore the actual application of this methods and tools. Cancer, stroke and Alzheimer's disease will be the three areas where informative examples will be collected, as they are complex conditions (many biological and environmental factors involved) and are representative of different approaches and degree of success in personalised medicine.

The main research questions addressed by the scoping review will be:

- What are the approaches to define the optimal size of stratification/validation cohorts?
- What are the differences, pros and cons of the prospective and retrospective nature of stratification and validation cohorts?
- What are the prerequisites and methods used for integration of multiple retrospective cohorts?

- Which validation designs exist for the stratification (or clustering) in personalised medicine? Which methods and tools are used to build the cohorts of validation (external/sub-cohort)? What are their gaps?
- What are the methods for the evaluation of the risk of bias?
- How are the (-omics, imaging, exposome, lifestyle etc.) data generated?
- What are the tools used for data management and multimodal data analysis used in personalised medicine (for instance, Galaxy)? What are their gaps?
- What quality of data of cohorts is needed to obtain a biomarker or multimodal data profiling? Are there requirements to monitor the collection of associated clinical data?
- What is the outlook of data generation seen as (CE-labelled) in-vitro diagnostics?

2. Methods for machine learning applied to stratification

The scoping review will focus on:

- Supervised and unsupervised machine learning methods for biomedical stratification using omics data. Few examples for other data modalities, e.g. imaging data, digital pathology and mobile sensor data will also be explored but not as the major focus.
- Cover both disease-based stratification (patient omics clustering, major focus) and drug-based stratification (clustering of drug-induced changes in patient-derived cells, minor focus)
- Methodologies that have been successfully validated and applied in clinical practice. New emerging approaches, which have not yet been sufficiently validated will also be explored but not as the major focus.
- Pros/cons, opportunities/limitations of different stratification methodologies and the associated validation approaches.
- Examples of successful applications.

Methodologies that have led to clinically validated biomarker signatures will be prioritised, as well as methodologies that have been cross-validated and externally tested on large sample sizes (preferably across multiple patient cohorts). Methods that lack statistical validation and a demonstrated biomedical application will be excluded.

The main research questions addressed by the scoping review will be:

- What are the main types of supervised and unsupervised machine learning methods for omics-based stratification in biomedicine (structured categorization)?
- What are the used and recommended workflows for supervised and unsupervised omics-based stratification (pre-processing, quality control, model building, model validation, model interpretation)?
- What are the specific strengths/weaknesses and opportunities/limitations of different types of omics-based stratification methods?
- Which validation methods exist for omics-based stratification models (assess accuracy, confirm biomedical relevance, test robustness) and what are their pros and cons?
- Which practical utility has been demonstrated for omics-based stratification and validation methods in real-world biomedical applications in the past (representative examples for previous success and/or failure stories, lessons learned)?
- What are the current gaps in standardization and methodological guidelines, and what is the outlook for the future of the field of omics-based machine learning stratification (new emerging approaches, new initiatives for data sharing, quality improvement, FAIRification)?

3. Pre-clinical methods for translational development of stratified therapies and treatments selection

The scoping review will focus on two aspects:

3.1. Personalised clinical decision-making based on pre-clinical models, aimed to explore drug sensitivity screening step (cellular based assay, organoids, PDX model) to predict therapy response and allocation of patients to different treatment arm, dose ranges and other aspects relevant for initiation of clinical trials. Suitable use cases will be selected in fields other than oncology, where clinical trials have been performed using pre-clinical models for personalised clinical decision-making.

The main research questions addressed will be:

- What are the fields of medicine other than oncology where pre-clinical models for personalised clinical decision-making have been applied?
- What are the pre-clinical models preferentially used in this context?
- How many drugs have been developed/are currently under development based on multi-omics profiling programs? What is the estimated success rate of the trial using this approach?
- What are the current gaps for broad implementation of pre-clinical testing for treatment selection?
- What information was collected at the pre-clinical stage to inform the clinical study design?

3.2. Stratified medicine development, to show which pre-clinical models (cellular, animal, organoid, in silico) are currently used as validation methods prior to personalised medicine clinical trials, both in academia and in industry. The example use case will be oncology.

As prospects, the review will discuss how to adapt the existing pre-clinical model systems to personalised medicine, and emerging models (such as in silico) which can replace the traditional animal models (3Rs). We will also perform a categorisation based on relevance and interpretation of models in the context of personalised medicine.

The main research questions addressed will be:

- Which pre-clinical models are currently used to provide validity data prior to therapeutic clinical trials of personalised medicine in oncology?
- What are the pros and cons of the various pre-clinical methods?
- Are the current pre-clinical models predictive for personalised medicine trials in oncology?

4. Methods for clinical trials in personalised medicine

The scoping review will focus on:

- Clinical trials, especially randomised trial designs, for personalised medicine.
- Trials evaluating a treatment in a subgroup of patients defined e.g. by a biomarker, in several clusters or subgroups of patients (e.g., basket or umbrella trials), trials comparing a personalised medicine strategy to a non-personalised strategy, or trials aiming at defining a subgroup of patients with enhanced response to treatment (e.g., adaptive enrichment design, adaptive signature design).

- Elements of clinical trial design applied to personalised medicine improving their appropriateness for HTA decision (e.g., external validity, choice of comparator, use of clinically meaningful outcome measure).
- Methodological reports (e.g., a scientific piece of work aiming at describing and evaluating the operational characteristics of a particular design) and guidance documents issued by regulatory or agencies for health technology assessment.
- Examples of published or ongoing trials in personalised medicine.

The review will not be restricted to a given medical field, although several examples in oncology are expected.

The main research questions addressed by the scoping review will be:

- What are the available designs for clinical trials applied to personalised medicine?
- What are the examples of current applications of these approaches?
- What are the pros and cons of the different approaches?
- What are the gaps in the current research on personalised medicine clinical trials?
- How is a personalised medicine strategy vs. non-personalised strategy evaluated?

Study identification

Relevant studies and documents will be identified balancing feasibility with breadth and comprehensiveness of searches.

Formal literature searches will be conducted on relevant databases (i.e., Medline, Embase, Cochrane Library) by the methods team. The keywords for the search strategy will be defined with the support of the review teams. Additional rounds of literature searches may be needed to refine specific aspects. The reference list of all identified reports and articles will be searched for additional studies.

To identify reports not published as scientific journal papers and unpublished (grey literature) information each review team will hand searching of relevant literature and websites (including conference meetings). Review teams may also contact relevant stakeholders to retrieve additional studies.

Documents published between 2005-2020 written in English, French, Spanish, Italian, German will be sought. Other specific time window, if deemed necessary by each review team, will be applied. Appropriate and clear justification for choices will be provided.

Appendix 1 reports examples of the search strategies planned for the four parts of this scoping review.

Eligibility Criteria

Each review team defined broad eligibility criteria based on the four “Scope & Research Questions” sections.

1. Methods for stratification and validation cohorts

We will include articles and other reports describing the methods applied to cohorts that have been used for patient stratification or validation of patient clustering obtained through stratification cohorts.

We will also include reports on methods to define the optimal size of cohorts, to design these cohorts, to integrate multiple retrospective cohorts, to evaluate risk of bias, and to manage data and analysis in personalized medicine. We are also interested in the quality of data and monitoring

of associated clinical data requirements and in the legal framework of data generated in personalized medicine.

Three case models will be explored: oncology, Alzheimer's disease and stroke.

These three fields were chosen for their big impact on society and individual health, because they are in three different phases of personalized medicine, which allows us to know different methods and strategies in different levels of development, and because they cover different kind of data to stratify patients. Oncology is the field where personalised medicine was firstly applied and where targeted therapies and diagnostics have been focused. Moreover, several applications of biomarkers for the successful stratification of patients with a given type of cancer exist, most of them based on molecular data, specially genomics. Alzheimer's disease research in personalized therapies and diagnostics is nowadays giving its firsts results, based in imaging, cognitive and also molecular data. Stroke is currently opening up to personalized medicine, with some approaches and studies in more initial steps. Most of the data for patient's stratification are imaging and molecular data. The review will cover a broad range of multimodal data profiling studies and biomarkers based on all kinds of data: genetic, metabolomic, genetic expression, genomic, or radiomic.

As general approach, we will search for (systematic) reviews to first identify the most common methodological approaches. Subsequent rounds of more specific searches will be conducted according to the results obtained from the scan of the reviews and to cover detailed aspects.

2. Methods for machine learning applied to stratification

We will include articles and other reports with methodology descriptions or reviews/opinion articles on supervised and unsupervised machine learning approaches and associated validation methods for omics-based stratification that have been tested on real-world biomedical data.

We will prioritize reports describing methodologies that have led to clinically validated biomarker signatures and those describing methodologies that have been cross-validated and externally tested on large sample sizes (preferably across multiple patient cohorts)

Articles reporting on methods that lack appropriate validation statistics and a demonstrated biomedical application will be excluded

There will be no restrictions in terms of types of publication or medical areas.

3. Pre-clinical methods for translational development of stratified therapies and treatments selection

3.1. Pre-clinical models for personalised clinical decision-making.

We will include articles and other reports describing methods (i.e. cellular based assay, organoids, animal models) used to assign treatment options to patient clusters. The case model will be mental disorders disease, chosen as non-oncology medical field. Indeed, this therapeutic area is included in the FDA Table of Pharmacogenomic Biomarkers in Drug Labelling as one of the most represented after oncology (9). Biomarkers in the table include but are not limited to germline or somatic gene variants (polymorphisms, mutations), functional deficiencies with a genetic etiology, gene expression differences, and chromosomal abnormalities; selected protein biomarkers that are used to select treatments for patients are also included.

3.2. Stratified medicine development in oncology

We will include articles and other reports describing translational medical approach, specifically pre-clinical validation methods applied prior to personalised medicine clinical trials. The case

model will be oncology, chosen as the field where personalised medicine was firstly applied and where targeted therapies and diagnostics have, for the most part, been focused.

The review will have a broad focus on the preclinical methodologies used for personalised medicine i.e. animal (mainly PDX), organoid, cellular models and in silico/computerised models, assessing the validity, reliability and predictive value of the various models. As general approach, we will include papers which describe the concept of the methods and exclude those which only deal with models applied to a specific type of cancer and original biomarker research.

Subsequent rounds of more specific searches will be conducted if needed, according to the results obtained from the scan of the first set of articles to cover detailed aspects.

There will be no restrictions in terms of types of publications included.

4. Methods for clinical trials in personalised medicine

We will include methodological and statistical articles and reviews describing or evaluating designs and validation of randomised controlled trials for personalised medicine, assessing both pharmaceutical and non-pharmaceutical interventions. We will also include articles reporting on personalised medicine trials and trial protocols, either published or available on trial registries. Finally, guidance documents issued by regulatory or health technology assessment agencies will be assessed.

There will be no restrictions in terms of types of publication or medical areas.

Study selection

The title and abstracts of records identified by the literature search will be screened by two independent reviewers. The full text publication of relevant articles will be retrieved and checked for confirming eligibility. Discrepancies will be solved by discussion among the review team and the method group if needed. An iterative approach to study selection is expected: each major change from what is reported in this protocol will be recorded and justified.

The screening process will be summarised in flow diagrams as suggested by the PRISMA guidelines for reporting scoping review (10).

Data extraction

The main feature of each report considered eligible, as providing information of a given aspect covered by one or more research questions, will be summarised in tables by one reviewer and checked by a second to ensure data quality. As we expect the reviews to include a variety of scientific articles and other documents, we will not develop a common pre-defined extraction form. However, the following information will be sought and summarised for each included report. This list will be adapted according to the needs of the different review teams.

- Author(s)/reference/title
- Year of publication
- Source origin/country of origin
- Type of publication (e.g. article, editorial, report, poster, etc.)
- Concept/Aims/purpose
- Study population and sample size (if applicable)
- Methodology/Study design
- Intervention type and comparator (if applicable)
- Duration of the intervention/time horizon (if applicable)
- Outcome measures (if applicable)

- Main results/findings
- Key findings that relate to the review question

This list will be adapted according to the needs of the different review teams.

Study quality

As general approach, we will not perform a formal assessment of methodological quality of the included studies as it is generally not performed in scoping reviews. However, the evaluation of risk of bias of clinical studies included as case examples may be considered.

Plan for presenting the results

The collected evidence will be assembled, summarized and reported to address the research questions defined in the Scope & Research Questions section. The format will be refined toward the end of the process when we will have the increased awareness of the contents of their included studies. Results will be discussed considering the gaps in methodology and the implications for policy, practice and research to inform the consultation exercise.

Consultation exercise

The activities of the review teams (WP3-WP6 in the PERMIT project, permit-eu.org/) will cover this aspect, through dedicated consultations and workshops with field experts. The discussion will involve PERMIT participants and associated partners, and the PERMIT project Scientific Advisory Board.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 874825.

References

1. Personalized medicine coalition http://www.personalizedmedicinecoalition.org (Accessed Feb 19, 2020).
2. Schleidgen S. et al. What is personalized medicine: sharpening a vague term based on a systematic literature review. BMC Medical Ethics 2013;14:55.
3. De Grandis G and Halgunset V. Conceptual and terminological confusion around personalised medicine: a coping strategy BMC Medical Ethics 2016;17:43.
4. Advice for 2018 2020 of the Horizon 2020 Advisory Group for Societal Challenge 1 "Health, Demographic Change and Well-being"". https://ec.europa.eu/research/health/pdf/ag_advice_report_2018-2020.pdf (Accessed Feb 19, 2020).
5. Di Paolo A. et al. Personalized medicine in Europe: not yet personal enough? BMC Health Services Research 2017;17:289.
6. UK Medical Research Council Stratified medicine methodology framework. https://mrc.ukri.org/research/initiatives/precision-medicine/stratified-medicine-methodology-framework/ (Accessed Feb 19, 2020).
7. Peters MDJ. et al. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthcare 2015;13:141-146.
8. Tricco AC. et al. A scoping review on the conduct and reporting of scoping reviews. BMC Medical Research Methodology 2016;16:15.
9. FDA. Table of Pharmacogenomic Biomarkers in Drug Labeling https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling (Accessed April 22, 2020)
10. Equator Network. Enhancing the QUAlity and Transparency Of health Research https://www.equator-network.org/reporting-guidelines/prisma-scr/ (Accessed Feb 19, 2020).

Box 1: Main definitions of personalised medicine

Proponent	Definition	Reference
Horizon 2020 Advisory Group and Strategic Research and Innovation Agenda (SRIA) of PerMed	Personalised medicine is 'a medical model using characterisation of individuals' phenotypes and genotypes (e.g. molecular profiling, medical imaging, lifestyle data) for tailoring the right therapeutic strategy for the right person at the right time, and/or to determine the predisposition to disease and/or to deliver timely and targeted prevention	https://ec.europa.eu/info/research-and-innovation/research-area/health-research-and-innovation/personalised-medicine_en http://www.permed2020.eu
European Council conclusions on personalised medicine for patients (2015/C 421/03)	Medical model using characterisation of individuals' phenotypes and genotypes, or tailoring the right therapeutic strategy for the right person at the right time, and to determine the predisposition to disease and/or deliver timely and targeted prevention.	https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015XG1217(01)&from=EN
UK Medical Research Council	Stratified medicine is the identification of key sub-groups of patients within a heterogeneous disease population; these being distinguishable groups with differing mechanisms, risk or course of disease, or particular responses to treatments. Stratification can be used to: • Improve mechanistic understanding of disease processes and enable the identification of new targets for treatments • Develop biomarkers for disease risk, diagnosis, prognosis and response to treatment • Allow treatments to be developed, tested and applied in the most appropriate patient groups	https://mrc.ukri.org/research/initiatives/precision-medicine/stratified-medicine-methodology-framework/
Personalized Medicine Coalition (PMC)	Personalized medicine is an evolving field in which physicians use diagnostic tests to determine which medical treatments will work best for each patient. By combining the data from those tests with an individual's medical history, circumstances and values, health care providers can develop targeted treatment and prevention plans. Personalized medicine is the tailoring of medical treatment to the individual characteristics of each patient. The approach relies on scientific breakthroughs in our understanding of how a person's unique molecular and genetic profile makes them susceptible to certain diseases. This same research is increasing our ability to predict which medical treatments will be safe and effective for each patient, and which ones will not be. Personalized medicine may be considered an extension of traditional approaches to understanding and treating disease. Equipped with tools that are more precise, physicians can select a therapy or treatment protocol based on a patient's molecular profile that may not only minimize	http://www.personalizedmedicinecoalition.org/, http://www.personalizedmedicinecoalition.org/Userfiles/PMC-Corporate/file/pmc_age_of_pmc_factsheet.pdf

	harmful side effects and ensure a more successful outcome, but can also help contain costs compared with a “trial-and-error” approach to disease treatment	
Precision Medicine Initiative (US NIH)	Precision medicine is "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person." This approach will allow doctors and researchers to predict more accurately, which treatment and prevention strategies for a particular disease will work in which groups of people. It is in contrast to a one-size-fits-all approach, in which disease treatment and prevention strategies are developed for the average person, with less consideration for the differences between individuals.	https://ghr.nlm.nih.gov/primer/precisionmedicine/definition
Food and Drug Administration	Precision medicine, sometimes known as "personalized medicine" is an innovative approach to tailoring disease prevention and treatment that takes into account differences in people's genes, environments, and lifestyles. The goal of precision medicine is to target the right treatments to the right patients at the right time.	https://www.fda.gov/medical-devices/vitro-diagnostics/precision-medicine
Schleidgen et al.	Personalized medicine seeks to improve stratification and timing of health care by utilizing biological information and biomarkers on the level of molecular disease pathways, genetics, proteomics as well as metabolomics.	BMC Medical Ethics 2013, 14:55
Sadée and Dai	Pharmacogenomics is a harbinger of personalised medicine, a paradigm shift from the mindset of ‘one-drug-fits-all’ to ‘the right drug for the right patient at the right dose and time.’ This does not mean that each patient will be treated differently from every other patient, an economically untenable proposition. Rather, patients are divided into groups by genetic and other markers that predict disease progression and treatment outcome.	Human Molecular Genetics 2005;14(suppl_2):R207–R214

Appendix 1: Examples of search strategies

1. Methods for stratification and validation cohorts

	Query	Results
#24	#21 AND #22 AND ([english]/lim OR [french]/lim OR [german]/lim OR [italian]/lim OR [spanish]/lim)	**429**
#23	#21 AND #22	**438**
#22	[embase]/lim NOT [medline]/lim	**9595932**
#21	(#18 OR #19) AND [2005-2020]/py	**1302**
#20	#18 OR #19	**1355**
#19	#17 AND [review]/lim	**779**
#18	#4 AND #9 AND #13 AND #16 AND ([cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim)	**818**
#17	#4 AND #9 AND #13 AND #16	**20180**
#16	#14 OR #15	**5241925**
#15	cancer*:ti,ab OR carcinoma*:ti,ab OR tumor*:ti,ab OR tumour*:ti,ab OR oncolo*:ti,ab OR leukemia*:ti,ab OR lymphoma*:ti,ab OR sarcoma*:ti,ab	**4465029**
#14	'neoplasm'/exp/mj	**3703629**
#13	#10 OR #11 OR #12	**29597086**
#12	validation:ti,ab OR method*:ti,ab	**9431867**
#11	'procedures'/exp	**28915500**
#10	'validation study'/exp	**81960**
#9	#5 OR #6 OR #7 OR #8	**1592114**
#8	'prospective study'/exp	**588853**
#7	'cross-sectional study'/exp	**339070**
#6	'cohort analysis'/exp	**560295**
#5	'cohort studies':ti,ab OR 'cohort study':ti,ab OR 'cohorts design':ti,ab OR 'prospective cohort':ti,ab OR 'retrospective cohort':ti,ab OR 'data integration':ti,ab OR bias:ti,ab OR 'cross study':ti,ab OR 'cross studies':ti,ab	**511111**
#4	#1 OR #2 OR #3	**535142**
#3	'personalized medicine'/exp	**41477**
#2	'biological marker'/exp	**296253**
#1	'stratified medicine':ti,ab OR biomarker*:ti,ab OR 'precision medicine':ti,ab OR 'personalized medicine':ti,ab OR 'personalised medicine':ti,ab OR 'individualized medicine':ti,ab OR 'individualised medicine':ti,ab OR 'individualized therapy':ti,ab OR 'individualised therapy':ti,ab OR 'patient stratification':ti,ab OR pharmacogenetics:ti,ab OR 'patient specific modeling':ti,ab OR 'personalized clinical decision making':ti,ab OR 'personalised clinical decision making':ti,ab OR 'prediction of response':ti,ab OR 'prediction of responses':ti,ab	**420990**

2. Methods for machine learning applied to stratification

Query		
#25	#23 AND #24	**688**
#24	[embase]/lim NOT [medline]/lim	**9568801**
#23	#20 AND #21 AND ([english]/lim OR [french]/lim OR [italian]/lim OR [spanish]/lim)	**1423**
#22	#20 AND #21	**1433**
#21	omic*:ti,ab OR 'machine learning':ti,ab OR 'personalized medicine':ti,ab OR 'personalised medicine':ti,ab	**59092**
#20	#4 AND #10 AND #16 AND #19	**4830**
#19	#17 OR #18	**6287177**
#18	validation:ti,ab OR validity:ti,ab OR validated:ti,ab OR 'cross validation':ti,ab OR 'cross validated':ti,ab OR test*:ti,ab OR 'clinical utility*':ti,ab OR accuracy:ti,ab OR robustness:ti,ab OR reliability*:ti,ab OR sensitivity:ti,ab OR specificity:ti,ab OR benchmark*:ti,ab OR bias:ti,ab OR 'cross study:ti,ab' OR 'cross studies':ti,ab	**6150811**
#17	'validation study'/exp OR 'reliability'/exp OR 'sensitivity and specificity'/exp OR 'benchmarking'/exp	**580344**
#16	#14 OR #15	**1174400**
#15	omic*:ti,ab OR 'omic based':ti,ab OR 'multi omic*':ti,ab OR genomic*:ti,ab OR transcriptomic*:ti,ab OR proteomic*:ti,ab OR metabolomic*:ti,ab OR lipidomic*:ti,ab OR epigenomic*:ti,ab OR microarray:ti,ab OR 'rna seq':ti,ab OR 'mass spectrometr*':ti,ab	**852339**
#14	#11 OR #12 OR #13	**758269**
#13	'mass spectrometry'/exp	**455591**
#12	'microarray analysis'/exp	**68369**
#11	'omics'/exp OR 'genomics'/exp OR 'epigenetics'/exp	**299009**
#10	#5 OR #6 OR #7 OR #8 OR #9	**5000844**
#9	'individualized medicine':ti,ab OR 'individualised medicine':ti,ab OR 'individualized therapy':ti,ab OR 'individualised therapy':ti,ab	**3459**
#8	'personalised medicine':ti,ab	**1713**
#7	'personalized medicine':ti,ab	**13669**
#6	'stratified medicine':ti,ab OR cluster*:ti,ab OR 'sub group*':ti,ab OR subgroup*:ti,ab OR biomarker*:ti,ab OR diagnos*:ti,ab OR prognos*:ti,ab OR 'precision medicine':ti,ab	**4904827**
#5	'biological marker'/exp OR 'personalized medicine'/exp	**330768**
#4	#1 OR #2 OR #3	**200079**
#3	'machine learning'/exp	**193633**
#2	'statistical learning'/exp	**46**
#1	'machine learning':ti,ab OR 'statistical learning':ti,ab OR 'supervised learning':ti,ab OR 'unsupervised learning':ti,ab	**34557**

3. Pre-clinical methods for translational development of stratified therapies and treatments selection

#15	#13 AND #14	**680**
#14	[embase]/lim NOT [medline]/lim	**9583919**
#13	#3 AND #6 AND #9 AND #12	**1133**
#12	#10 OR #11	**2126059**
#11	'disorders of higher cerebral function'/exp/mj OR 'mental disease'/exp/mj OR 'psychosis'/exp/mj	**1448175**
#10	'psychiatric disease*':ti,ab OR 'mental disorder*':ti,ab OR 'psychiatric disorder*':ti,ab OR 'mental illness':ti,ab OR depression:ti,ab OR 'bipolar disorder':ti,ab OR bipolarism:ti,ab OR anxiety:ti,ab OR 'personality disorder*':ti,ab OR 'psychotic disorder*':ti,ab OR schizophreni*:ti,ab OR 'eating disorder*':ti,ab OR 'trauma related disorder*':ti,ab OR 'post traumatic stress disorder*':ti,ab OR 'substance abuse disorder*':ti,ab OR 'asperger syndrome':ti,ab OR autism:ti,ab OR 'delirium tremens':ti,ab OR epilep*:ti,ab OR 'hallucinogen related disorder*':ti,ab OR hysteria:ti,ab OR 'minor depressive disorder*':ti,ab OR 'major depressive disorder*':ti,ab OR 'obsessive compulsive disorder*':ti,ab OR 'obsessive compulsive personality disorder*':ti,ab OR 'schizoaffective disorder*':ti,ab OR 'schizoid personality disorder*':ti,ab OR alzheimer:ti,ab OR dementia:ti,ab	**1372445**
#9	#7 OR #8	**6083613**
#8	'drug therapy'/mj	**243267**
#7	'therapy selection':ti,ab OR 'therapeutic selection':ti,ab OR 'treatment':ti,ab OR 'patient allocation':ti,ab OR 'drug therapy':ti,ab OR 'trial success rate':ti,ab OR 'therapy selected':ti,ab OR 'therapeutic selected':ti,ab	**5925767**
#6	#4 OR #5	**433005**
#5	'biological marker'/exp/mj OR 'personalized medicine'/exp/mj	**93445**
#4	'stratified medicine':ti,ab OR biomarker*:ti,ab OR 'precision medicine':ti,ab OR 'personalized medicine':ti,ab OR 'personalised medicine':ti,ab OR 'individualized medicine':ti,ab OR 'individualised medicine':ti,ab OR 'individualized therapy':ti,ab OR 'individualised therapy':ti,ab OR 'patient stratification':ti,ab OR 'patient specific modeling':ti,ab OR 'personalized clinical decision making':ti,ab OR 'personalised clinical decision making':ti,ab OR 'prediction of response':ti,ab	**412990**
#3	#1 OR #2	**680024**
#2	'animal model'/exp/mj OR 'organoid'/exp/mj OR 'in vitro study'/exp/mj OR 'translational research'/exp/mj OR 'disease model'/exp/mj OR 'cell culture'/exp/mj OR 'preclinical study'/exp/mj	**236722**
#1	'cellular model*':ti,ab OR 'drug development*':ti,ab OR 'drug response':ti,ab OR 'drug evaluation':ti,ab OR 'drug evaluated':ti,ab OR 'patient specific modeling':ti,ab OR organoid*:ti,ab OR 'in silico':ti,ab OR 'drug response assay':ti,ab OR 'drug sensitivity screening':ti,ab OR 'pdx models':ti,ab OR 'patient derived xenografts':ti,ab OR 'preclinical pdx':ti,ab OR 'humanised mouse model':ti,ab OR 'preclinical model*':ti,ab OR 'pre clinical stage' OR 'pre clinical testing':ti,ab OR 'translational medical research':ti,ab OR 'disease model*':ti,ab OR 'translational model*':ti,ab OR 'animal model*':ti,ab OR xenograft*:ti,ab	**474356**

4. Methods for clinical trials in personalised medicine

No.	Query	Results
#14	#11 AND #12 AND ([english]/lim OR [french]/lim OR [german]/lim OR [italian]/lim OR [spanish]/lim)	**927**
#13	#11 AND #12	**929**
#12	[embase]/lim NOT [medline]/lim	**9610086**
#11	#7 OR #10	**1221**
#10	#4 AND #5 AND #8 AND [2020-2020]/py	**202**
#9	#4 AND #5 AND #8	**7669**
#8	'clinical trial*':ti,ab	**514125**
#7	#3 AND #4 AND #5 AND [2005-2020]/py	**1026**
#6	#3 AND #4 AND #5	**1033**
#5	design*:ti,ab OR methods:ti OR method:ti,ab	**4793126**
#4	'biological marker'/exp/mj OR 'personalized medicine'/exp/mj OR 'stratified medicine':ti,ab OR biomarker*:ti,ab OR 'precision medicine':ti,ab OR 'personalized medicine':ti,ab OR 'personalised medicine':ti,ab OR 'individualized medicine':ti,ab OR 'individualised medicine':ti,ab OR 'individualized therapy':ti,ab OR 'individualised therapy':ti,ab	**431819**
#3	#1 OR #2	**52941**
#2	'clinical trial'/exp/mj	**50652**
#1	'basket trial*':ti,ab OR 'basket stud*':ti,ab OR 'multi arm*':ti,ab OR 'master protocol*':ti,ab OR 'platform stud*':ti,ab OR 'platform trial*':ti,ab OR 'umbrella trial*':ti,ab OR 'adaptive stud*':ti,ab OR 'adaptive trial*':ti,ab OR 'umbrella stud*':ti,ab	**2402**