Health Canada Santé Canada

Draft guidance document on the collection and analysis of disaggregated data in clinical trials

Canada

Foreword

Guidance documents provide assistance to industry and health care professionals on how to comply with governing statutes and regulations. They also provide guidance to Health Canada staff on how mandates and objectives should be met fairly, consistently and effectively.

Guidance documents are administrative, not legal, instruments. This means that flexibility can be applied. However, to be acceptable, alternate approaches to the principles and practices described in this document must be supported by adequate justification. They should be discussed in advance with the relevant program area to avoid the possible finding that applicable statutory or regulatory requirements have not been met.

Health Canada reserves the right to request information or material, or define conditions not specifically described in this document, to help us adequately assess the safety, efficacy or quality of a therapeutic product. We are committed to ensuring that such requests are justifiable and that decisions are clearly documented.

This document should be read in conjunction with the accompanying notice and the relevant sections of the *Food and Drug Act* and its *Regulations* and other applicable guidance documents.

Table of contents

Overview

Background

This guidance will help sponsors interpret amendments to the *Food and Drug Regulations* (regulations) requiring them to submit clinical trial data disaggregated by sex, age and race/ethnicity if they have already done so as required by existing legislation or regulation in the United States or Europe.

It is standard practice to collect disaggregated subgroup data during drug development (for example, refer to the international standards outlined ICH M4E(R2)).

This guidance:

- sets out Health Canada's expectations
- supports our efforts to be transparent
- ensures that the provisions in the regulations on disaggregated data are interpreted consistently

The first part of this guidance document outlines how to comply with regulatory requirements for submitting disaggregated data. The second part of this guidance provides information and best practices related to collecting, analyzing and reporting on disaggregated data in clinical evidence, where applicable.

Disaggregated data is data obtained by breaking down aggregated datasets into subcategories. These subcategories are based on factors such as sex, age, race/ethnicity, co-morbidities and/or other patient demographics.

Disaggregating data is an important approach to data collection and analysis in drug development, as it:

- allows sponsors and Health Canada to assess the efficacy and safety of drug products in different subgroups
- ensures there is consistency with the overall results, where feasible
- indicates where more focused post-market monitoring may be needed to further verify safety and efficacy in certain populations after products are approved for sale in Canada

The collection and analysis of disaggregated data is at the core of Health Canada's sex and gender-based analysis plus (SGBA Plus) mandate.

SGBA Plus is an analytical process used to assess:

- how diverse groups of women, men and non-binary people may experience policies, programs and initiatives
- differences in how groups of people (subgroups) react to a specific drug in terms of both efficacy and safety

The "plus" in SGBA Plus acknowledges that SGBA goes beyond biological (sex) and socio-cultural (gender) differences. It indicates the importance of intersectionality, which refers to the multiple factors that intersect to make us who we are, such as:

- sex
- age
- race
- gender
- geographic location

Policy objective

Clinical trial participants should represent the population who will be using the drug. This is important as intrinsic (genetic, physiologic) and extrinsic (cultural, environmental) characteristics of a population, including age, sex and race/ethnicity, can lead to different treatment effects. Some, but not all, of the intrinsic and extrinsic factors that can impact a person's response to a drug product are known.

Patient populations are also often heterogeneous across many factors (for example, sex, age and race).

These factors highlight the importance of recruiting representative clinical trial participants so that data can be disaggregated and analyzed meaningfully by subgroups of interest (for example, sex, age, race, co-morbidities, and medication use).

Health Canada is leveraging experience in other jurisdictions by aligning with existing legislation and regulations of the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA).

Learn more about intrinsic and extrinsic differences in the following guidance and studies:

- ICH E5: Ethnic factors in the acceptability of foreign data (ICH Harmonised Tripartite Guideline, PDF)
- ICH E8: General considerations for clinical studies (PDF)

Policy statement

The amendments to the regulations are intended to:

- receive disaggregated data and subgroup analyses that can help assess differences in key efficacy and/or safety parameters between clinically relevant, prognostic subgroups
- increase transparency by assessing and reporting on the diversity of clinical trial participants for each drug product

Scope and application

This draft guidance document is for sponsors who will be applying to Health Canada for market authorization for a drug for human use.

The scope of this guidance document includes drugs for human use. These drugs are regulated by the Health Products and Food Branch in accordance with the *Food and Drugs Act* and the regulations.

Terms and definitions

Disaggregated data: Data obtained by breaking down large-scale or aggregated datasets into subcategories, such as sex, age and race/ethnicity, or a combination of these (and/or other) subcategories.

Disaggregated data plan: A plan designed to support the collection and assessment of disaggregated data by age, sex, racial/ethnic groups and associated covariates. The plan is used to establish whether and how a treatment may affect the benefit-risk profile of specific subgroups.

Gender: The socially constructed roles, behaviours, expressions and identities of girls, women, boys, men and gender-diverse people. It influences how people perceive themselves and each other, how they act and interact, and the distribution of power and resources in society.

Gender identity is not confined to a binary (girl/woman, boy/man) or static. It exists along a continuum and can change over time.

There is considerable diversity in:

- how individuals and groups understand, experience and express gender through the roles they take on
- the expectations placed on them
- a person's relations with others
- the complex ways that gender is institutionalized in society

Please refer to the following definition by the Canadian Institute of Health Research:

- What is gender? What is sex?

Heterogeneity assessment: An evaluation examining whether differences in treatment efficacy or safety are anticipated between subgroups of the overall study population. This assessment is a key aspect of clinical trial protocol design. It should describe the measures taken to establish whether there are, or could be, clinically important age, sex and racial/ethnic differences in response to study treatment.

Sex: A set of biological attributes in humans and animals. It is mainly associated with physical and physiological features, such as chromosomes, gene expression, hormone levels and function, and reproductive/sexual anatomy.

Sex is usually categorized as female or male, but there is variation in the biological attributes that comprise sex and how those attributes are expressed.

Sex differs from gender, which refers to the socially constructed roles, behaviours, expressions and identities of girls, women, boys, men and gender-diverse people.

Regulatory Requirements

Submitting disaggregated data

Health Canada now requires that new drug submissions (NDS) and Level I supplements to new drug submissions (SNDS) for human drugs include clinical evidence that has been disaggregated.

For information on these requirements, please consult:

- Post-notice of compliance changes: Safety and efficacy document

As a first step, these requirements will apply when sponsors:

- have submitted clinical data disaggregated by sex, age and race/ethnicity
    - based on legislative and/or regulatory requirements for market authorization set out by the U.S. Food and Drug Administration (FDA) or the European Medicines Agency (EMA)

We expect sponsors to conduct additional subgroup analyses when relevant (for example, co-morbidities, current and past medication use). This is standard practice in clinical trial design.

Currently:

- FDA requires data disaggregated by sex, race/ethnicity and age
- EMA requires data on sex and age, along with a justification if that data is not available

When both datasets are available, we prefer FDA data for sex, age and race/ ethnicity where possible.

When sponsors have not previously submitted disaggregated data to the FDA and/or EMA, we strongly encourage them to provide disaggregated data to Health Canada, where feasible. You are to provide a justification if your submission does not include this data. The justification should include the rationale for why disaggregated data are not applicable to the current submission, and/or why they have not been included.

In general, clinical trials are not designed to assess efficacy and safety in subgroups with statistical rigour. Nor do they have a sufficient number of participants to facilitate such an assessment. For these reasons, we understand that the results from such analyses are descriptive, exploratory in nature and should be interpreted with caution.

Disaggregated data best practices

About disaggregated data best practices

Recruiting diverse clinical trial participants and conducting relevant subgroup analyses remains critically important to ensure the possibility of determining:

- whether differences between subgroups exist based on all of the clinical evidence
- if further monitoring after the drug reaches the market is necessary (where feasible and applicable)
- whether the clinical trial population mirrors the full diversity of the population that will use the drug product being tested

The collection and analysis of disaggregated data is a key aspect in all phases of a health product's life cycle:

- study design
- dose selection
- testing and evaluation
- clinical validation
- risk analysis and management
- transparency and reporting
- post-market safety monitoring

Considerations in protocol design

This section outlines information to consider in clinical trial protocol design for drugs for human use, especially Phase 3 and 4 trials. Such information will help sponsors recruit diverse clinical trial participants and plan for appropriate subgroup analyses.

As a standard practice, the diversity of clinical trial participants should reflect the population that will use the drug. Before an application is submitted for market authorization in Canada, drugs should be evaluated in participants who represent the full range of persons likely to receive the product, whenever possible.

For example:

- Considerations for inclusion of women in clinical trials and analysis of sex differences

It is also important that clinical trial results are applicable to the populations likely to use the product. Evidence that can help evaluate benefit-risk profiles of specific subgroups is generated by collecting:

- disaggregated data on the age, sex and race/ethnicity of participants
- other clinically relevant baseline characteristics

As mentioned, differences in age, sex and race/ethnicity may contribute to different efficacy and safety profiles for drug products, for example:

- sex differences in pharmacokinetics and pharmacodynamics may impact drug safety and effectiveness in some contexts

Similarly, data from diverse racial/ethnic groups may necessitate different treatment parameters, for example:

- different dosing recommendations for therapeutics, such as:
    - lower initial starting dose of statins in Asian patients or
    - higher doses of tacrolimus in Black patients

It is understood that racial/ethnic differences in drug metabolism are largely attributable to genetic and genomic differences. These differences can directly influence drug absorption, distribution, metabolism and excretion.

Genetic and genomic factors influence the overall enzyme make-up of an individual, and certain mutations or gene variants are more common in specific ethnic groups. This contributes to pharmacokinetic variability in these subgroups. For example, gene variants affecting transporter proteins impact the safety and efficacy of statins. Similarly, gene variants also affect expression of enzymes, which influence the bioavailability and clearance of tacrolimus.

For more information:

- The pharmacogenomics of statins
- Safety and efficacy of statins in Asians
- Sex differences in pharmacokinetics and pharmacodynamics
- Genotype-guided tacrolimus dosing in African American kidney transplant recipients
- Impact of ABCG2 and SLCO1B1 polymorphisms on pharmacokinetics of rosuvastatin, atorvastatin, and simvastatin acid in Caucasian and Asian subjects: A class effect?
- African-American race modifies the influence of tacrolimus concentrations on acute rejection and toxicity in kidney transplant recipients

Heterogeneity assessment

A heterogeneity assessment (HA) is an evaluation examining whether differences in treatment efficacy or safety are anticipated between subgroups of the overall study population. The HA should describe the measures taken to establish whether there are, or could be, clinically important age, sex and racial/ethnic differences in response to the drug product being tested.

The U.S. Food and Drug Administration (FDA) outlines a similar approach in the following:

- Diversity plans to improve enrollment of participants from underrepresented racial and ethnic populations in clinical trials: Draft guidance for industry

To ensure diverse subgroups are meaningfully considered in a drug product's development and testing, an overarching HA should be performed at the start of the clinical development program. This HA will determine how best to approach clinical development and trial design.

Sponsors should update the HA as needed throughout the development process as more information on sources of heterogeneity is learned from early phase trials. These additional sources of heterogeneity may require further supporting data.

If the HA indicates that treatment differences are not anticipated among subgroups, enrollment should still be broad and reflect disease epidemiology. Broad enrollment of representative populations:

- permits subgroup analyses
- may enable the assessment of effect consistency across trials, where appropriate

Learn more:

- ICH E9: Statistical principles for clinical trials

Alternatively, if the HA or evidence from previous or ongoing studies indicates meaningful differences in any of the subgroups studied, sponsors should consider such differences when planning, designing and analyzing individual studies within the clinical development program. This includes plans to recruit enough

representative patients to be able to analyze the benefit-risk profile in the relevant subgroups, where possible.

Learn more in the following EMA guidance document:

- Guideline on the investigation of subgroups in confirmatory clinical trials

If there is a possibility that there could be differing treatment responses between subgroups, sponsors should explore the potential reasons for the heterogeneity and describe this in the trial(s) design(s) stage. Analyses should consider these potential differences. Note: In some cases, this type of analysis may only be feasible when analyzing all of the clinical evidence and/or in post-market monitoring.

Known factors that may affect the safety and efficacy of the drug product between subgroups should inform the HA. Sponsors should consider disease pathophysiology, previously identified predictive or prognostic factors, biological plausibility, epidemiology of the condition and/or therapeutic drug class.

Information for certain diseases may be limited, making it difficult to assess potential subgroup-specific heterogeneity. To supplement their HA, sponsors should find relevant data sources in academic literature and/or seek out real world evidence (RWE).

Sponsors may use non-clinical studies to provide supportive data. For example, pharmacology and toxicology studies conducted in male and female animals can indicate potential sex-related differences in concentration response, safety or efficacy.

To understand how non-clinical studies can support and inform clinical trial evidence, consult ICH harmonized guidelines for clinical trials, such as:

- ICH E8 : General considerations for clinical studies
- ICH S9 : Nonclinical evaluation for anticancer pharmaceuticals
- ICH S11 : Nonclinical safety testing in support of development of paediatric pharmaceuticals
- ICH M3(R2) : Nonclinical safety studies for the conduct of human clinical trials for pharmaceuticals

As the HA informs the study design and enrollment plan, sponsors should include it in their submission material for clinical trial applications and/or market authorization.

Of note, the contents of the current guidance apply to preceding documents such as:

- Guidance document for clinical trial sponsors: Clinical trial applications

Disaggregated data plan

A disaggregated data plan (DDP) is used to determine the parameters for disaggregated data collection and subsequent analyses. By supporting the collection and assessment of disaggregated data by age, sex, racial/ethnic groups and other associated covariates, the DDP helps sponsors establish if and how a treatment may affect the benefit-risk profile of specific subgroups.

The content of the DDP should be based on the findings of the HA. To support subgroup analysis, sponsors should aim to enroll representative proportions of age, sex and racial/ethnic subgroups that are consistent with those who will use the drug product or as indicated by the HA.

Otherwise, sponsors should either:

- justify how the enrollment criteria provide acceptable representation for the intended population or
- explain if a factor other than age, sex or race/ethnicity is a primary driver of heterogeneity upon which the disaggregated data plan should be based

These recommendations generally apply to all trials, but more feasibly, are targeted to phase 3 and 4 studies, as these trials tend to enroll more patients than earlier trials.

The DDP and data submission should describe:

- the number of participants planned for each study
    - outline how the sponsor will collect subgroup data throughout the product life cycle in the description
- how age, sex and race/ethnicity, as well as other related covariates, will be assessed
- how potential differences will be taken into account
    - for example, by enriched or restricted enrollment (by selecting for or limiting certain patient characteristics), confirmatory trials and so on, as applicable

This description will maximize early discussions of subgroups and establish a strategy for considering diverse subgroups before the trials begin, if required. For example:

- if it is known early on that the treatment is only effective in certain subgroups, then trial enrollment could be restricted to those populations
- if a factor affecting subgroups is known to be prognostic in value, then treatment groups may be balanced for this factor at the randomization stage through stratification

Sponsors should capture the details of both the HA and DDP, including enrollment strategies, in the appropriate sections of the clinical study protocol and statistical analysis plan (SAP). Include applicable results, analyses and discussion in the clinical study report (CSR).

For data submitted in the CSR, sponsors should:

- capture disaggregated data for enrolled and completed participants
    - population pharmacokinetic analyses often include age and sex as co-variates
    - race/ethnicity should also be included as it may affect the interpretation of drug product efficacy and/or safety in subgroups as well as clinical parameters of use
    - pharmacometric modelling and analyses are encouraged
- present key efficacy and safety data by age, sex and race/ethnicity and describe whether any treatment modifications (for example, of dose or dosing regimen) are needed for subgroups
    - clinically relevant findings may be reflected in the product labelling as per Health Canada's guidance on the product monograph

We encourage sponsors to engage in discussions with the relevant Health Canada directorate early on in the process. Sponsors may also ask for feedback during pre-clinical trial application meetings and/or in pre-submission meetings for a new drug submission (NDS) or supplement to a new drug submission (SNDS).

More details on pre-submission meetings for NDS or SNDS are included in the following guidance document:

- Management of drug submissions and applications

Data collection and analysis

Important considerations for collecting and analyzing disaggregated data include key subgroups and parameters for analysis, interpretation and reporting.

Sponsors should pre-specify the strategy for enrolling and assessing subgroup data in the protocol before they enroll clinical trial participants.

Key subgroups

The variables that should be considered when designing clinical trials are specific to the particular therapeutic area in which the study is being conducted. However, there are key considerations for including clinical trial participants who are diverse in sex, age and race/ethnicity. Of note, it is understood that representative enrollment of age, sex, and racial/ethnic subgroups may pose challenges.

Sponsors should consider potential reasons if they have difficulty recruiting certain populations. In this case, you should devise strategies to enhance enrollment and retention in the subgroups of interest, as needed.

Learn more by consulting the following guidance from the U.S. Food and Drug Administration (FDA):

- Enhancing the diversity of clinical trial populations - Eligibility criteria, enrollment practices and trial designs

Sex

Include a representative number of females and males in clinical trials, as applicable to the intended population that will use the drug.

To identify and account for potential sex-related differences when planning clinical trials, we recommend that females (such as those who are of child-bearing age or menopausal) be included as early as possible in the clinical trial research stage.

For details on what to consider when collecting and analyzing sex-related data, please consult the following guidance document:

- Considerations for inclusion of women in clinical trials and analysis of sex differences

Age

Age groups that reflect the clinical prevalence of disease should be adequately represented. This permits the analysis of various age groups, as needed.

We recommend that detailed disaggregation of age groups be provided. For example, to better inform treatment decisions for geriatric patients in clinical practice, older patients may be grouped into ages 65 to 74, 75 to 84 and over 85 rather than the combining all age groups into a 65+ category.

Learn more in the following guidance documents:

- ICH E7: Studies in support of special populations: Geriatrics
- FDA: Inclusion of older adults in cancer clinical trials

Detailed disaggregation of pediatric age subgroups corresponding to developmental biology specific to drug product pharmacology is also recommended, for example:

- preterm newborn infants
- term newborn infants (0 to 27 days)
- infants and toddlers (28 days to 23 months)
- children (2 to 11 years)
- adolescents (12 to 16/18 years)

Neurocognitive development, puberty, hormonal changes and weight are some examples of processes or characteristics that should be considered when planning, analyzing and reporting pediatric subgroup data.

Learn more:

- ICH E11: Clinical investigation of medicinal products in the pediatric population

Race/ethnicity

To assess potential race/ethnicity-related differences for the population who will use the drug, we recommend that representative proportions of subjects be enrolled in clinical trials. Note that the following will vary:

- specific race/ethnicity subgroups depending on the population that is expected to use the drug
- the geopolitical context where the clinical trials will be held and where the application for market authorization will be submitted

For trials conducted in Canada, self-reported race/ethnicity should follow Statistics Canada classifications. Current classifications for population groups include:

- Arab
- Black
- White
- Filipino
- Korean
- Chinese
- Japanese
- West Asian
- South Asian
- Latin American
- Southeast Asian
- other group (specify)

Indigenous Peoples comprising First Nations, Métis and Inuit populations are another key population group in Canada. Respondents should be able to select the categories that apply.

For trials conducted elsewhere, sponsors should follow the reporting categories of the respective national regulatory body, if available. More detailed racial/ethnic data may be required in certain conditions with specific hereditary patterns.

Reporting standards in different countries will vary, making it difficult to pool data for comparisons. Consequently, sponsors should clearly define subgroups in their protocols.

Analyzing, interpreting and reporting disaggregated data

Sponsors should stratify results for key efficacy and safety outcome(s) by age, sex and race/ethnicity, as applicable. These results should be interpreted with caution, given the potential for a limited number of participants in some subgroups and for confounding by other factors.

For guidance on analyzing and interpreting subgroup data, please refer to:

- Section 5.7, "Subgroups, interactions and covariates" of ICHE9

If the differences observed among various subgroups are interpretable and considered clinically relevant, sponsors should investigate if the results are biologically plausible or may have occurred by chance. Following a thorough assessment, sponsors should address those results that are biologically plausible and thus may impact the benefit-risk profile of the product in the relevant subgroup(s).

In such cases, Health Canada may:

- request additional data
- implement terms and conditions
- recommend specific labelling
- require additional post-market studies or surveillance
- require additional risk minimization measures

Sponsors should include:

- a description of methods and approach in the clinical trial protocol and statistical analysis plan (SAP), as applicable
- results, analyses and subsequent discussion in the clinical study report (CSR)

To assess the consistency of the efficacy and safety findings across relevant subgroups, sponsors should present results based on analyses described in the disaggregated data plan (DDP) (and captured in the appropriate sections of the clinical study protocol and SAP). These should be placed in the appropriate section of the CSR, if available.

In particular, sponsors should also include:

- a table that presents disaggregated demographic data by treatment group (ICH M4)
- subgroup results for key measures in the corresponding efficacy and safety sections, as applicable
- a discussion of differences that are considered clinically relevant and biologically plausible on the benefit-risk profile specific to age, sex or race/ethnicity, as applicable

Transparency

Health Canada will continue to communicate up-to-date information about drugs for human use. You can find the following information online:

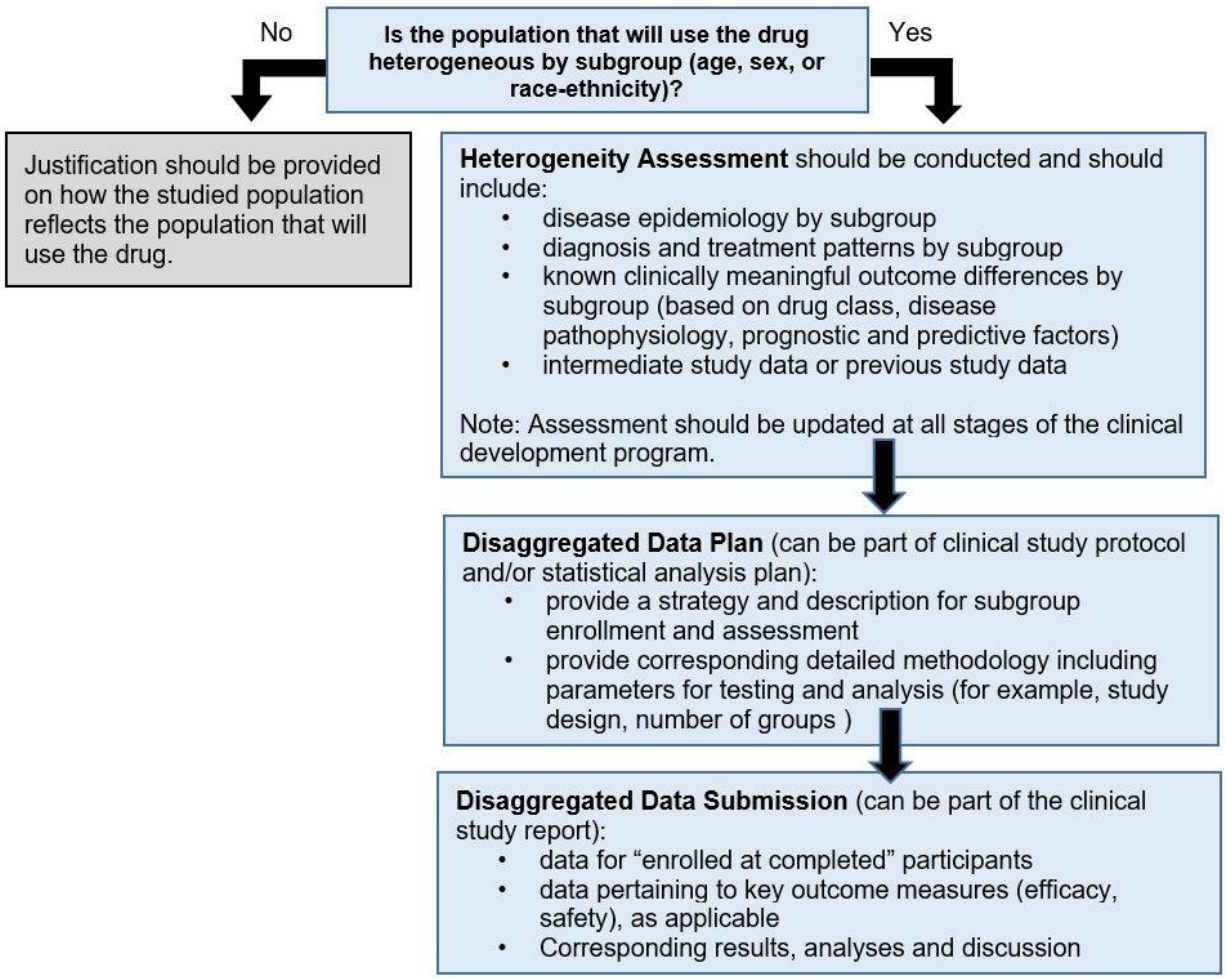
- Notice of Compliance (NOC) database: contains NOCs issued for drugs for human and veterinary use
- Drug product database: contains information about drug identification numbers (DINs) issued for drugs for human and veterinary use, including the product monograph for human drugs and the product labelling for veterinary drugs
- Drug and Health Product Portal: contains Regulatory Decision Summaries and Summary Basis of Decision documents, which describe Health Canada's rationale for the approval of prescription drugs for human use
- Clinical information portal: contains the clinical information, including disaggregated data, filed by sponsors to seek approval of human drugs under Division 8 of the *Food and Drug Regulations* (FDR)

Information posted will not contain any confidential business information.

Disaggregated data decision tree

The disaggregated data decision tree outlines key questions to guide clinical trial design, based on the population that will use the drug being tested. The decision tree summarizes the content in "Considerations in protocol design".

Figure 1. Decision tree: Study design



Contact us

Direct your questions to the appropriate directorate in Health Canada:

Biologic and Radiopharmaceutical Drugs Directorate (BRDD)
Office of Regulatory Affairs
Health Products and Food Branch
Health Canada
Email: brdd.ora@hc-sc.gc.ca

Pharmaceutical Drugs Directorate (PDD)
Health Products and Food Branch
Health Canada
Email: pharma_drug_enquiries-renseignements_medicaments_pharma@hc-sc.gc.ca

Natural and Non-prescription Health Products Directorate (NNHPD)
Health Products and Food Branch
Health Canada
Email: nnhpd.consultation-dpsnso@hc-sc.gc.ca

Marketed Health Products Directorate (MHPD)
Health Products and Food Branch
Health Canada
Email: mhpd-dpsc@hc-sc.gc.ca

For questions or comments on this guidance document:

Centre for Policy, Pediatrics and International Collaboration
Biologic and Radiopharmaceutical Drugs Directorate
Health Products and Food Branch
Health Canada
Email: brdd-cppic_brdd-cppci@hc-sc.gc.ca

Related links

Government of Canada

- What is gender-based analysis plus
- Guidance document: Considerations for inclusion of women in clinical trials and analysis of sex differences

FDA

- FDA disaggregated data requirements
- Guidance document: Diversity plans to improve the enrollment of participants from underrepresented racial and ethnic populations in clinical trials

EMA

- EMA disaggregated data requirements

ICH

- ICH guidelines in the safety and efficacy series