COMPARING MANUFACTURER SUBMITTED ANALYSIS AND COMMON DRUG REVIEW REANALYSIS OF RESULTS: A REVIEW OF HEALTH TECHNOLOGY ASSESSMENT REPORTS FOR NON-ONCOLOGY MEDICATIONS FROM 2018 TO 2022

By Fatemeh Mirzayeh Fashami

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree

Master of Science in Health Research Methodology

MASTER OF SCIENCE (2023)	McMaster University
HEALTH RESEARCH METHODOLOGY	Hamilton, Canada
TITLE:	COMPARING MANUFACTURER SUBMITTED ANALYSIS AND COMMON DRUG REVIEW REANALYSIS OF RESULTS: A REVIEW OF HEALTH TECHNOLOGY ASSESSMENT REPORTS FOR NON-ONCOLOGY MEDICATIONS FROM 2018 TO 2022
AUTHOR:	Fatemeh Mirzayeh Fashami
SUPERVISOR:	Dr. Mitchell Levine
SUPERVISORY COMMITTEE:	Dr. Jean-Eric Tarride Dr. Behnam Sadeghirad
NUMBER OF PAGES:	54

ABSTRACT

Introduction

Identifying key differences between the manufacturer's submitted analysis and the Canada's Common Drug Review (CDR) economic reanalysis is a crucial step toward creating more appropriate models by manufacturers. We compared manufacturers' submitted analysis to CDR reanalysis in order to identify any variations in incremental costs utility ratio (ICUR) and 3-year Budget impact analysis (BIA). We assessed the willingness to pay (WTP) threshold and CDR critiques on manufacturers' clinical and economic reports.

Method

A pair of reviewers extracted data regarding therapeutic category, percent price reduction requested by CDR, WTP, and the critiques on the manufacturers' clinical and economic reports in publicly available CDR reports from 2018 to 2022. We used Wilcoxon rank test to assess the difference between mean incremental QALY, ICUR, and BIA in manufacturers and CDR reanalysis reports and chi-square tests and logistic regression to assess the relationship between the variables and the final CDR recommendation.

Results

Of 178 reports assessed, 31 received "do not reimburse" recommendation and 147 received "reimburse with criteria or conditions". The median ICUR in manufacturer's analysis was $138,658/QALY and significantly lower than ICUR reanalyses by CDR of $380,251/QALY. The ICUR in manufacturers' submitted reports was 2.5-fold lower than in the CDR reanalysis

($138,658/QALY versus $380,251/QALY). The CDR reanalysis median for 3-year BIA was $4,575,102 which was 27% higher than the manufacturers submitted 3-year BIA (*p* value<0.001).

The most frequent CDR critiques were clinical effectiveness and the uncertainty of evidence in cost-effectiveness analysis and miscalculations in the population of patients and the percentage of market share in BIA.

From 2018 to 2020, $100,000 was the most frequent WTP threshold followed by a $50,000 threshold, but during 2021 and 2022, the CDR only used $50,000 as a WTP threshold.

Conclusion

Manufacturers may tend to underestimate the costs or overestimate the effect of their medications.

Table of Contents

List of tables

List of figures

Abbreviations

BIA	Budget Impact Analysis
CDR	Common Drug Review
DRD	Drug for Rare Disease
EQ-5D	EuroQol 5-Dimension
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HUI	Health Utilities Index
ICUR	Incremental Cost-Utility Ratio
OECD	Organisation for Economic Co-operation and Development
pCODR	Pan-Canadian Oncology Drug Review
pCPA	pan-Canadian Pharmaceutical Alliance
QALY	Quality Adjusted Life Years
SF-36	Short Form 36
CADTH	The Canadian Agency for Drugs and Technologies in Health
INESSS	The Institut national d'excellence en santé et en services sociaux
PMPRB	The Patented Medicine Prices Review Board
WTP	Willingness To Pay

Declaration of Academic Achievement

The work described in this thesis was performed Fatemeh Mirzayeh Fashami, who will be referred to as the "primary researcher" moving forward. This project was supervised by Dr. Mitchell Levine and the committee members were Dr. Jean-eric Tarride and Dr. Behnam Sadeghirad.

Kimia Hariri and Amirreza Peyrovinasab performed data extraction, under the guidance of the primary researcher who was responsible for the study design. Additionally, the primary researcher supervised and resolved disagreements in data extraction when it was needed. The primary researcher conducted data analysis and interpretation and wrote the manuscript.

Dr. Mitchell Levine supervised developing the concept, study design, analysis and checked all the outputs. Dr. Jean-Eric Tarride and Dr. Behnam Sadeghirad assisted in study design, analysis and interpretation.

1. Chapter one: Introduction and Background

1.1. Rationale

1.1.1. The Canadian Health care budget and its constraints

Canada is among the high spending countries with publicly funded healthcare and placed above the Organisation for Economic Co-operation and Development (OECD) average per-person healthcare expenditure. Total healthcare expenditures experienced an increase from $269 billion and 11.5% of the total Gross domestic product (GDP) in 2015, to $305 billion and 13.2% of the GDP in 2020 and it is projected to expend $331 billion in 2022. (1) From the total healthcare expenditures of 2022, it is projected that 24.3% will be spent at hospitals, 13.6% will be paid to physicians, and 13.6% will be allocated to medications. The cost of medications continues to rise due to the increasing demand for newer high-cost drugs, the aging population, innovation in new medications and personalized medicines. Medication cost in Canada increased for 1.6% and 4.1% in 2020 and 2021. Drug costs include all costs for the provincial and territorial governments, private insurances, and patients' out of pocket and are projected to continue to rise, reaching a 5.4% increase by 2022, with a total projected cost of $14.5 billion.(1)

Canada is recognized for its extensive healthcare system, which provides its citizens universal access to vital medical services. The government of Canada and the provincial and territorial governments implements various policies and programs to prevent excessive drug pricing to provide the budget allocations for medications, with the intention that citizens receive affordable and high-quality healthcare.(2) One of the important contributors to the escalation of healthcare expenditure is the demographic shift towards an aging population.(3) With increasing

longevity, there is a concomitant surge in medical requirements and treatments.(3) The elderly population tend to use healthcare services more frequently than younger age group.(3)

The Canadian healthcare system is currently facing financial constraints, with insufficient resources to address the escalating costs of healthcare provision. The government, having a limited budget allocated to the healthcare sector, must make choices in spending to preserve the quality and availability of healthcare for all citizens. With the demographic shift towards an aging population, the prevalence of chronic illnesses and the rising demand for healthcare services, and the Covid-19 pandemic, it has become increasingly difficult for the government to reconcile the requirement for quality care with the financial limitations imposed on the healthcare system. Therefore, the role of the CADTH in Canada excluding Quebec and the role of the Institut national d'excellence en santé et en services sociaux (INESSS) in Quebec is crucial in evaluating cost-effectiveness and budget impact analysis. Both agencies assess HTA reviews. However, the current study focuses on the CADTH non-oncology reviews of technology assessments in Canada outside of Quebec.

1.1.2. The Canadian Agency for Drugs and Technologies in Health (CADTH)

CADTH is an organization that assesses the economic evaluation, and it is shown that it had zero tolerance to accept medications without any clinical criteria or conditions with ICUR above the informal willingness to pay threshold. (4) Since 2010, the economic evaluation of oncology drugs has been assessed through pCODR process, while HTA for other medications has been assessed by the CDR process. Since 2015, CADTH has included drugs with evidence-based expanded use based on clinical effectiveness and safety data, even if Health Canada had yet to approve the medication for that indication. (5) In 2017, CADTH published the 4th edition of the

guideline for the Economic Evaluation of Health Technologies and added details on what manufacturers must consider in HTA reports, particularly in modeling, measurements, costs, and handling of uncertainty. This edition added BIA as a requirement for new submissions. (6) In addition, manufacturers are required to adhere to CADTH guideline and procedures when submitting a product.

1.1.3. CADTH reanalysis on the economic reviews

Upon receiving HTA reports from manufacturers, CDR and pCODR evaluate clinical evidence through a process that involves performing a literature review and collecting clinicians' and patients' opinions. In addition, CDR and pCODR assess the model structure and input data for non-oncology and oncology medicines. (6) After the internal process is completed by CDR and pCODR, CADTH publishes an economic report, including a critical appraisal of the manufacturer's report, a reanalysis with improved input data and model structure along with revised cost-utility results. Further, CADTH provides a non-binding recommendation for each submission to "Reimburse", "Do not reimburse", or "Reimburse with clinical criteria and/or conditions".

1.2. Identifying Relevant Literature

There have been a few studies published comparing the HTA reports submitted by manufacturers and CADTH reanalysis in both CDR and pCODR sections.

Rocchi 2012 investigated to identify the predictors associated with negative recommendations within the context of CDR reports from 2003 to 2009. CDR rejected 48% of the 138 reports analyzed, with a significant variance observed across different therapeutic categories, ranging from 0% for HIV to 88% for analgesic drugs. Clinical critiques for CDR rejection included inappropriate comparator, clinical uncertainty, and outcome identified as unacceptable.

Regression analysis identified four factors that significantly predicted the rejection of a drug: clinical factors, higher price compared to its comparators, request for reconsideration, and use of price as the only economic evidence. Using different $/QALY cut points showed that ICUR value cannot predict the final CDR recommendation.(7)

Rocchi 2018 analyzed data from 2010 to 2017 using the pan-Canadian Pharmaceutical Alliance (pCPA) archives to evaluate the level of agreement between the recommendations of CADTH and the negotiation decisions of the pCPA. The results of the study showed that the pCPA occasionally negotiated for products that CDR had recommended not to be listed. The study also found that the median ICUR for oncology products was more than double that of non-oncology products ($168K/QALY vs $70K/QALY). In addition, the recalculated ICURs by CADTH for CDR recommendations were significantly higher than those provided by the manufacturers and the ratio of recalculated ICUR to manufacturer ICUR was higher in CDR compared to the pCODR.(8)

Rocchi 2013 conducted an evaluation on the impact of using surrogate outcomes on the final CDR recommendations. Of the 156 studies between 2003 and 2010, 44% involved surrogate outcomes. The overall rejection rate for CDR requests was 48%, while for those involving surrogate outcomes, it was 41%. Non-accepted surrogates had a 68% rejection rate, while accepted surrogates had a 25% rejection rate. The authors concluded that the majority of surrogates were accepted by CDR, and that non-accepted surrogates were significantly associated with clinical uncertainty and a higher likelihood of rejection.(9)

Similar studies are available in pCODR published reports as well. Saluja 2021 assessed the difference between incremental costs, incremental QALYs, and ICURs of manufacturers' submitted results versus pCODR reanalysis for oncology drugs from 2012 to 2018. The proportion of ICURs

that were considered cost-effective from 2012 to 2018 were assessed to identify any probable trend. ICURs in pCODR reanalysis had a rising trend over time mainly due to the recalculation of time horizon and utility scores. The final recommendations were shown to be dependent on clinical evidence and patient preferences. Main methodologic flaws in manufacturer's submitted HTA were overestimation of time horizon, and utility scores followed by issues in costing (overestimation of cost calculations for comparator or other hospitalization costs or by not considering drug wastage), uncertainty due to indirect comparisons, flaws in model structure (bias in extrapolation in Partitioned Survival Models (PSMs), estimation of treatment benefits) and extrapolation of treatment benefit beyond treatment time. This study yielded mean incremental cost-effectiveness ratios (ICURs) of $134K/QALY for manufacturer-submitted analysis and the mean for lower and upper extremes for ICUR were $166K/QALY and $365K/QALY, respectively. The results showed that the manufacturers of oncology drugs may have a tendency to overestimate the cost-effectiveness of their medication.(10)

Masucci 2017 evaluated 39 economic evaluations of oncology drugs published on the CADTH website between 2011 and 2014. The authors identified several methodological flaws in these evaluations, including issues with time horizon, duration of treatment benefit, costing, utility estimation, model structure, extrapolation techniques, uncertainty in indirect comparison, calculation errors, and the quantity of clinical data available.(11)

Ball 2022 evaluated the HTA reports for oncology medicines in three countries: Canada, the UK, and Australia, between 2019 and 2020. Among the 32 reports from the pCODR the incremental QALY was 1.3 (0.13 to 4.34) in the manufacturer report, while the pCODR reanalysis showed a 60.3% reduction, with an incremental QALY of only 0.78 (0.08 to 2.25). Furthermore,

the incremental costs were USK$109K (USK$12K to USK$129K) in the manufacturer report and USK$201K (USK$41K to USK$984K) in the pCODR reanalysis, representing 2.8 times increase in costs. The ICUR value in the pCODR reanalysis was USD$201K/QALY, which was nearly double the manufacturer's ICUR value of USD$110K/QALY. The authors identified the most frequent criticisms in the reports, which included costing, time horizon, utility value, treatment benefit, extrapolation, comparator, and subgroup analysis.(12)

Nagase 2019 investigated all the CADTH reports for rare diseases from 2012 to 2018. The study analyzed 104 recommendations (42 CDR and 62 pCODR) and found a similar percentage of positive recommendations across various therapeutic classes of drugs. The study found that drugs deemed safe or provided benefits in clinical or patient-reported outcomes were more likely to receive positive recommendations for reimbursement. However, the study did not find any correlation between the type of recommendation and factors such as daily treatment cost, cost-effectiveness, or the condition being treated (whether it was cancer or non-cancer).(13)

1.3. Study rationale

In 2021 in Canada, 24.9% and 75.1% of drug costs were for oncology and non-oncology medications, respectively.(14) More than 75% of the medication budget was allocated to non-oncology therapies but more research evaluated pCODR HTA reports compared to CDR ones. While prior studies have investigated some methodological concerns, a formal comparison of manufacturer-calculated and CDR-reanalyzed ICURs in the Canadian setting has yet to be conducted since 2018. Furthermore, no studies have investigated the differences between the BIA calculated by manufacturers versus CDR or pCODR reanalysis, along with their respective critiques. The 4th edition of the CADTH guideline, which includes comprehensive reporting criteria, was

published in March 2017, and no study has assessed CDR recommendations since then.(6) Although non-oncology medications have a massive impact on the total drug budget of Canada, no studies have assessed the differences between manufacturer-submitted reports and CDR reanalysis reports and no studies evaluated economic appraisal by therapeutic categories, drugs for rare diseases, or unmet needs in Canada after 2018.

1.4. Overall Goal of the Study

This study examined whether manufacturer submitted HTA models generated lower ICURs and BIA than the CDR reanalysis models. This study's objectives were to measure the difference between manufacturer and CDR ICURs, incremental costs, incremental QALYs, and 3-year BIA and to evaluate whether the recommendation for reimbursement of drugs that addressed an unmet need or DRD differed from the overall recommendation rate. Additionally, the study explored recommendation rates across different drug categories. The study also examined the WTP threshold mentioned in various reports to determine if there was a trend in the WTP threshold of CDR reports and if WTP was different in DRD compared to other drugs. Furthermore, the study aimed to identify the main CDR critiques provided in each final recommendation.

1.5. Research question

Among all the CDR applications with a completed review from 01 January 2018 to 31 December 2022, the following questions were proposed:

1. What is the difference in mean economic values (incremental costs, incremental QALY, ICUR, and BIA) between manufacturer submitted HTAs and CDR reanalysis?
2. what is the difference between CDR recommendations for medicines that address an unmet need compared to other drugs?

3. what is the difference between recommendations for DRD and non-DRD?
4. what is the difference between recommendations for drugs in different therapeutic categories?
5. What are the WTP thresholds in CDR reports and what is the difference in the WPT thresholds of DRD compared to other medications?
6. What is the price reduction requested by CDR and if there is any significant difference between price reduction requested by CDR in drugs for DRD and non-DRD?
7. What are the methodological critiques (costs, effectiveness, modeling, and clinical evidence) in the cost-utility analysis of manufacturer submissions?
8. What are the methodological critiques in the BIA analysis of manufacturer submissions?

2. Chapter two: Methods

2.1. Selection of reports

All the CDR economic reports with final complete recommendations from 01 January 2018 to the 31 December 2022 were included. Since CADTH only publishes clinical and pharmacoeconomic reports after finishing its evaluation, we only included reports with "complete" status.

2.2. Data extraction procedure

A pair of reviewers extracted data independently in a standardized abstraction sheet using MS Excel. Data extractions were compared by two reviewers and disagreements were resolved through discussion and involvement of a third reviewers if needed.

The following information were extracted from CADTH website: project number, generic drug name, brand name, therapeutic area, CDR recommendation, and recommendation date. In addition, we extracted information on (1) study type (cost utility analysis vs cost minimization analysis), (2) model type (Markov, semi-Markov, partitioned survival model, decision tree, mixed model with repeated measures, or not reported), (3) data source (randomized controlled trials, observational studies, or combination of randomized trials and observational studies, and not reported), (4) utility measurement (EQ-5D, SF-36, HUI, Others, and not reported), and (5) if risk equation was used. We also extracted the information related to the outcomes like therapeutic category, addressing an unmet need, and DRD to investigate any potential association between these factors and the final recommendations. Additionally, we extracted both WTP and percentage price reduction requested by CDR.

2.3. Outcomes

We extracted data on incremental cost, incremental QALY, ICUR, and BIA for manufacturers' submissions and CDR's reanalysis. Since CADTH guideline requires manufacturers to provide comparisons to multiple medications, we extracted data for the lower and upper extremes of the ICURs and BIAs and then calculated the mean values. We excluded reports with dominant of dominated for one of the lower or upper extremes of ICUR. We extracted cumulative 3-year BIA reported by the manufacturer and BIA reanalyzed by CDR from CDR reports.

We extracted outcomes, including WTP, percentage of price reduction requested by CDR, and CDR critiques based on model structure, cost, utility score, effectiveness, and uncertainty levels.

The therapeutic category were based on the international classification of diseases, 11th revision (ICD-11) codes as (1) circulatory system diseases,(2) blood or blood-forming organs disease, (3) infectious and parasitic diseases, (4) musculoskeletal and connective tissue diseases, (5) immune system disease, (6) endocrine nutritional, and metabolic diseases, (7) mental and behavior, and neurodevelopmental disorders, (8) nervous system diseases, (9) respiratory system diseases, (10) sleep-wake disorders, (11) diseases of the visual system, (12) diseases of the ear or mastoid process, (13) diseases of the digestive system, (14) diseases of the skin, (15) diseases of the genitourinary system; (16) pregnancy, childbirth or (17) the puerperium and symptoms and signs or clinical findings, not elsewhere classified.

2.3.1. Contributions to any unmet needs

Including input from patient and clinician groups is a critical component of the clinical report prepared by the CDR to evaluate the presence of unmet needs for a given indication. After

a comprehensive review of the available clinical data, the CDR determines whether the data are sufficient to support the notion that a new intervention addresses an aforementioned unmet need. It is well-known that CADTH tends to recommend drugs with a price reduction request when drugs are priced beyond the willingness to pay threshold if an innovative therapy address an unmet therapeutic need, there is a lack alternatives, or it is a DRD.(10) Our study sourced data pertaining to the unmet need variable directly from the discussion in the final CDR recommendations.

2.3.2. Drug for rare disease

Our study has adopted the CADTH website's definition of DRD, which are drugs intended to diagnose, treat, prevent, or alleviate life-threatening, seriously debilitating, or both chronic and serious diseases or conditions. Specifically, these rare diseases or conditions are those that affect no more than 5 in every 10,000 individuals in Canada.(17)

2.3.3. Willingness to pay threshold.

We extracted explicit values for the WTP threshold, which were mentioned in the price reduction statements of the CDR recommendations. We assessed whether any changes in the WTP threshold for CDR had occurred over the 2018 to 2022 years.

2.3.4. CDR final recommendation

The CDR final recommendation to "reimburse with clinical criteria and/or conditions" is a highly general recommendation that fails to offer specific data regarding whether the criteria pertain to product's cost or clinical limitations. Consequently, we gathered additional data regarding the final recommendations and specified them within the following variables: if it specifically requires a prescription be generated by a specialized physician (based on medical

specialty or experience), if it imposes additional limitations on clinical criteria (e.g., limiting indications or requesting further trials for continued medication after a designated period), or if it requires a price reduction.

2.3.5. Critiques on the manufacturers' submitted reports

Based on the previous studies (7,11,13,15) and discussion with authors, we decided to use the critiques mentioned in the literature and add other critiques that CDR frequently uses in its reports.

2.3.5.1. Critiques on the model structure

We analyzed the most frequently recurring critiques of the CDR model structure to assess criticisms regarding model structure. The critiques were categorized into "Model does not reflect current practice", "Model structure did not capture some important aspects", "Lack of face validity or CADTH could not validated the model", "Lack of transparency", "Model was not flexible enough for CADTH to change some aspects", "Programming error", "Model is based on the treatment response instead of health states", "Other structure critiques" and "Any structure critiques".

2.3.5.2. Critiques on costs

We have categorized critiques related to costing into "Unacceptable extra costs (like hospitalization, administration, or other charges)", "Missing or underestimation of some costs", "Inappropriate list prices for the comparator(s)", "Other cost critiques" and "Any cost critiques".

2.3.5.3. Critiques on utility scores

In the fourth edition of the Canadian guidelines for the economic evaluation of health technologies specifies that utility scores should represent the health states in the model.(6) Hence,

compliance with the guideline and the necessity to revise the utility score in the manufacturers' reports were also included as variables and critiques on "the utility score(s) not representing health states" was assessed in CDR critiques. Furthermore, we extracted the presence of critiques relating to "Utility score(s) do not reflect Canadian ones", "Other utility critiques" and "Any utility critiques".

2.3.5.4. Critiques on effectiveness

We used the following list for effectiveness critiques: "Critiques on uncertainty of evidence", "Not enough clinical evidence available", "Clinical data uncertainty beyond clinical evidence", "Not a proper population captured", "Limitation in ITC or NMA", "Inappropriate comparator was used in the base case analysis", "AE not captured correctly", "Incorrect or double counting mortality", "Evidence with low number of participants", "Only using surrogate outcome", "Critiques on extrapolation", "Inappropriate cut off points", "Using interim data instead of final ones", "Critiques on risk equation", "Other effectiveness critiques" and "Any effectiveness critiques". We assessed if risk of bias and GRADE certainty of evidence was addressed and, when applicable, extracted risk of bias levels and if a specific tool was mentioned to measure risk od bias.

2.3.5.5. Critiques on BIA

BIA guideline of Canada (2020) states that critical points of calculating BIA should be listed as real target population and its growth over time, opting for a suitable comparator, cost consideration, sensitivity analysis, and characterizing uncertainty. (16) According to this guideline, the target population should be based on the approved drug indication and its limitation of usage if there is any. In addition, growth of the market should be calculated based on the anticipated

growth of the patient's target population. (16) We extracted data on BIA critiques, including "Population of patients", "Percent of patients covered by public plan", "Market share", "Treatment duration ", "Drug dosing", "Drug wastage or multiple administration from a vial", "Faults on drug discontinuation", "Intervention cost or intervention related costs", "Comparator cost or comparator related costs", "Not considering one/some comparators", "Adherence/compliance" or "Any other critiques due to the clinical data uncertainty" and "Other critiques".

2.3.5.6 Uncertainty level

We extracted level of uncertainty mentioned in economic reports' conclusions of the CDR reports since high uncertainty level could be a potential risk for rejection of new drugs.

2.6. Data management and analysis

The mean economic values (mean incremental costs, mean incremental QALY, mean ICUR, and mean 3-year BIA) in the manufacturers' base case reports and in the CDR's base case reanalysis were summarized using descriptive statistics. Wilcoxon rank test was used to evaluate if there were statistically significant differences between the medians in the manufacturer's reports versus the CDR's reanalysis reports with respect to mean incremental costs, mean incremental QALY, mean ICUR and mean BIA. Distribution plots were created for ICURs and 3-year BIAs in order to compare the manufacturer's reported values with the CDR reanalysis results.

The general characteristics of the studies, the number of drugs and the positive recommendation rate for rare disease, drugs addressing an unmet need, drug in different therapeutic categories, and drugs that were clinical superior (versus clinical non-inferior) were reported descriptively. In order to identify potential predictors, each variable in the database

underwent individual testing for independence against the categorical outcome. This was achieved using chi-squared tests, where the degrees of freedom were calculated as (rows - 1) * (columns - 1), and the significance level (a) was set at 0.05. Variables that exhibited a significant relationship with the outcome ($p < 0.10$) were considered dependent and were subsequently included in the multivariate regression model.

All statistical analyses were carried out using SPSS (IBM SPSS version 26 release 26.0.0.0, Armonk, NY).

3. Chapter three: Results

There were 178 CDR reports disseminated from 01 January 2018 to 31 December 2022. At the time of this evaluation, all reports had "complete" status in CADTH website, and their related clinical and economic reviews were published completely. A total of 31 reports received "do not reimburse" recommendation. Table S1 provides reasons for the rejection of these studies. Of the 178 reports analyzed, 147 (83%) received "recommended by clinical criteria and/or conditions," a phrase used by CADTH to indicate acceptance based on clinical conditions, price reduction, or both. Of these 147 reports, 128 reports (87%) were recommended for any clinical conditions and 138 (95%) were recommended for price reduction. Among the reports limiting recommendation to a clinical condition, 91 reports were recommended to be prescribed by a specialist or an experienced physician. The other limiting conditions included (1) reimbursement for a limited time, (2) a condition for conducting follow-up tests to measure treatment response, and (3) being prescribed only in a hospital setting. These limitations applied to 119 reports.

3.1. General Characteristics of reports

The majority of the reports were cost-utility analysis (87.1%), followed by cost comparison (7.3%), and cost-minimization (5.6%). Among all reports, 62.9% involved randomized controlled trials along with observational studies, while 37.1% used only randomized controlled trials. Markov or semi-Markov models were used in 74.8% of the models, followed by a combination of decision tree and Markov model (13.2%), decision tree only (6.6%), and others (5.3%). The data for general characteristics of the studies are demonstrated in table 1.

Results of the chi-square test showed that Study type (p value=0.793), trial characteristics (p value=0.159), model type (p value=0.429), and utility measurement tools (p value=0.181) are

independent of the final CDR recommendation and these variables cannot predict the final recommendations.

Table 1. General characteristics of reports

General submission characteristics	Final recommendation				Total	
	Reimburse with clinical criteria/ condition		Do not reimburse			
	N	%	N	%	N	%
Study type						
Cost-utility analysis	127	86.4	28	90.3	155	87.1
Cost-minimization analysis	9	6.1	1	3.2	10	5.6
Cost comparison	11	7.5	2	6.5	13	7.3
Trial characteristics						
Only Randomized Clinical Trials	51	34.7	15	48.4	66	37.1
Randomized Clinical Trials and Observational studies	96	65.3	16	51.6	112	62.9
Model type						
Markov or semi-Markov	94	75.8	19	70.4	113	74.8
Combination of Decision tree & Markov	16	12.9	4	14.8	20	13.2
Decision tree	6	4.8	4	14.8	10	6.6
Others	8	6.5	0	0.0	8	5.3
Utility measurement tools						
EQ-5D	93	71.0	21	72.4	114	78.1
SF-36	11	8.4	0	0	11	7.5
HUI	4	2.7	0	0	4	2.7
Others	14	10.7	3	10.3	17	11.6

"Accepted with clinical criteria/condition" is the explicit phrase that CDR use and it implies any restrictions in price or the clinical use

3.2. CDR recommendations based on therapeutic categories.

Among therapeutic categories, almost half the requests were from "Endocrine, nutritional, or metabolic diseases" (35 reports, 20%), "Diseases of the nervous system" (30 reports, 17%), and "Diseases of the respiratory system" (18 reports, 10%). Table 2 presents details on total number of reports, number of reports with reimbursement recommendations based on clinical criteria or

price reduction, number of reports requiring price reduction, and the total number of reports were recommended with price reduction in each drug category.

Results of chi-square test demonstrated that therapeutic category is not associated with final CDR recommendation (*p* value=0.319), and it is not a predictor of the final recommendation.

Table 2. Reimbursement recommendation rate in different therapeutic categories.

Item	Therapeutic category	Reimbursement with clinical criteria or condition "		Reimbursement with clinical criteria		Reimbursement with price reduction		Total reports	
		N	%	N	%	N	%	N	%
	Total reports	**147**	**83**	**128**	**72**	**138**	**78**	**178**	**100**
1	Endocrine, nutritional or metabolic diseases	28	80	27	77	27	77	35	20
2	Diseases of the nervous system	26	87	23	77	25	83	30	17
3	Diseases of the respiratory system	17	94	14	78	16	89	18	10
4	Certain infectious or parasitic diseases	14	93	8	53	13	87	15	8
5	Diseases of the digestive system	10	83	9	75	9	75	12	7
6	Diseases of the skin	8	67	8	67	8	67	12	7
7	Diseases of the visual system	8	73	6	55	8	73	11	6
8	Diseases of the musculoskeletal system or connective tissue	9	100	8	89	9	100	9	5
9	Diseases of the blood or blood-forming organs	6	75	6	75	6	75	8	4
10	Diseases of the immune system	7	88	7	88	7	88	8	4
11	Mental, behavioural or neurodevelopmental disorders	4	67	3	50	4	67	6	3
12	Diseases of the circulatory system	4	67	4	67	2	33	6	3
13	Diseases of the genitourinary system	3	100	2	67	3	100	3	2
14	Symptoms, signs or clinical findings, not elsewhere classified	1	50	1	50	0	0	2	1
15	Sleep-wake disorders	0	0	0	0	0	0	1	1
16	Diseases of the ear or mastoid process	1	100	1	100	0	0	1	1
17	Pregnancy, childbirth or the puerperium	1	100	1	100	1	100	1	1

"Accepted with clinical criteria/condition" is the explicit phrase that CDR use and it implies any restrictions in price or the clinical use

3.3. Price reduction requested based on the therapeutic category

CDR requested a median price reduction of 63.5% (interquartile range: 35.0% to 87.75%). The price reduction requested was lower for drugs in "Certain infectious or parasitic diseases" category, followed by "Diseases of the skin", and "Diseases of the musculoskeletal system or connective tissue." The highest median CDR requested price reduction was for "Diseases of the blood or blood-forming organs", "Diseases of the respiratory system," and "Endocrine, nutritional or metabolic diseases". Data for the other categories are presented in Figure 1.

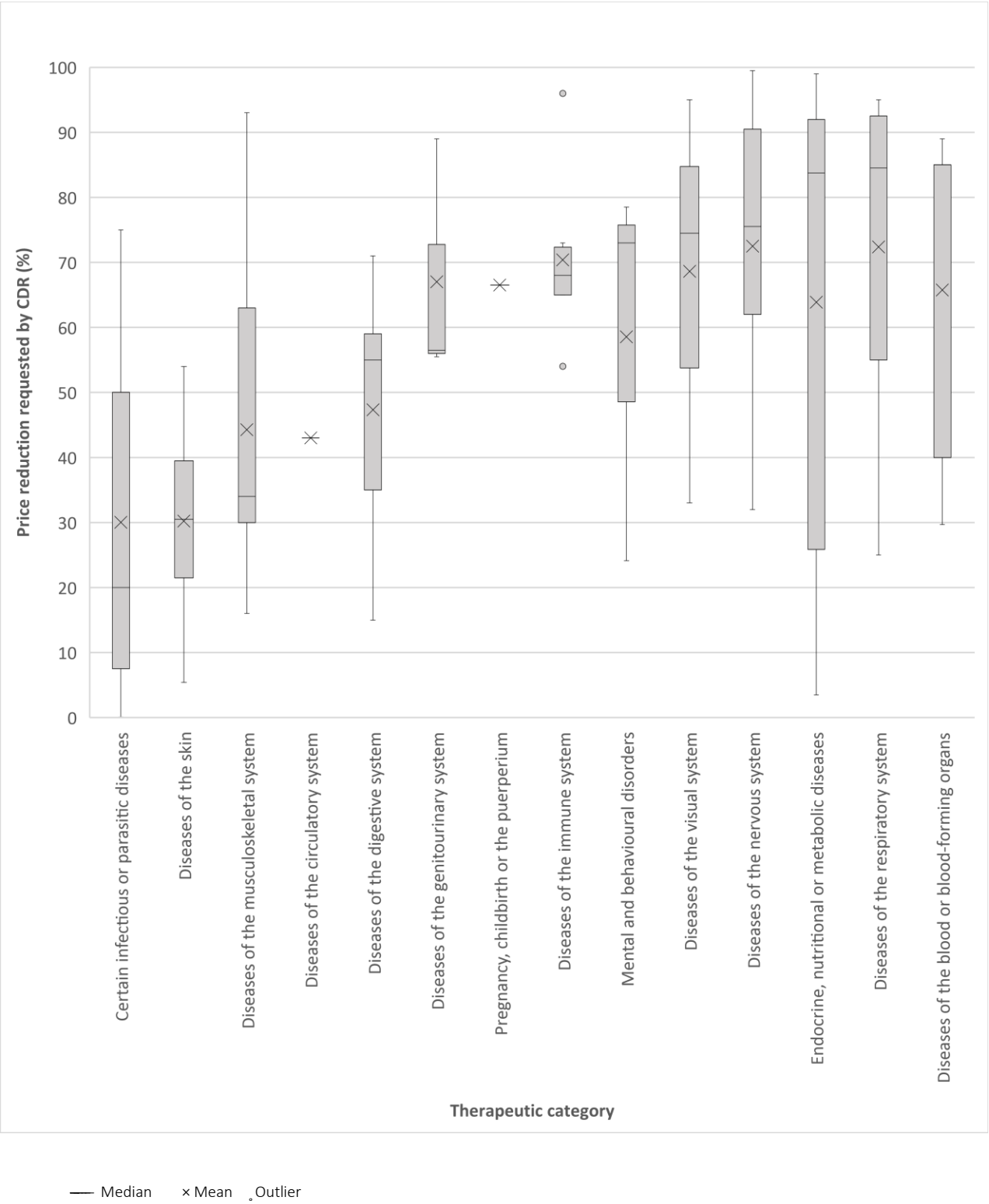

Figure 1. Price reduction requested by CDR based on therapeutic categories

3.4. Price reduction requested based on clinical superiority and non-inferiority

Of the 178 reports, in 62 (34.8%) of the CDR reports stated that the new intervention is superior to the currently available medications and accepted the cost-utility model. For many medications, manufacturers submitted cost-utility analysis, but CDR considered them non-inferior interventions and consequently assumed zero incremental QALY for them. CDR rejected the new interventions as clinically superior products in these cases and provided a cost-minimization analysis instead. Among 155 cost-utility reports, CDR considered 93 (60%) of them as non-inferior medications. The price reduction requested by CDR was higher for medicines with clinical superiority (median of 80.7%, interquartile interval 57.7% to 93.0%) compared to the one that were non-inferior to the current medications (median of 55.0%, interquartile interval 29.7% to 67.5%). The related data is demonstrated in Figure 2.

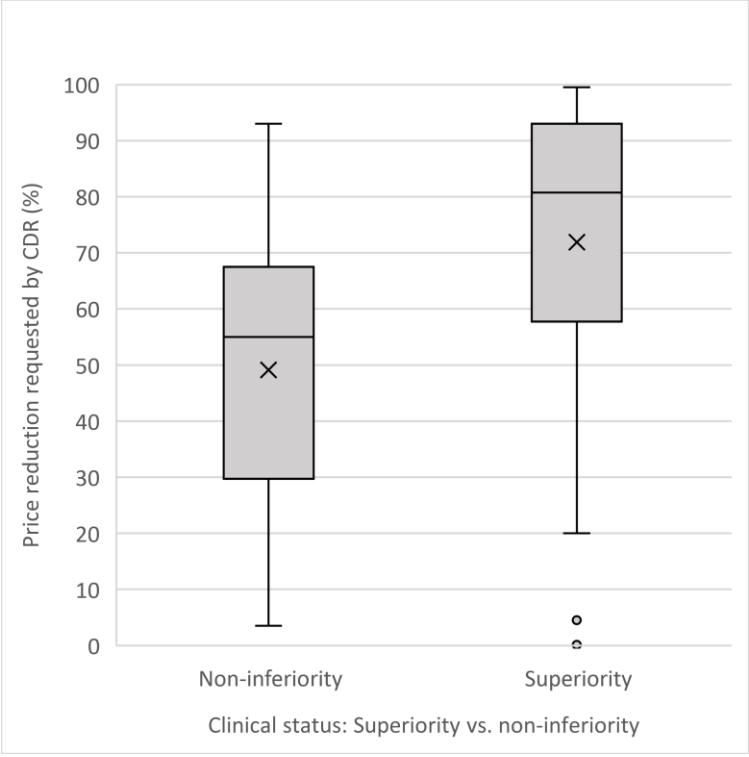

Figure 2. Price reduction requested by CDR based on clinical superiority or non-inferiority

3.5. Price reduction requested based on DRD and non-DRD

The median percent price reduction that was requested for DRD was 87.0% (interquartile interval 65.0% to 95.0%) and was less for non DRD 55.0% (interquartile interval 30.0% to 73.0%). Figure 3 demonstrates the price reduction requested for DRD vs. non-DRD.

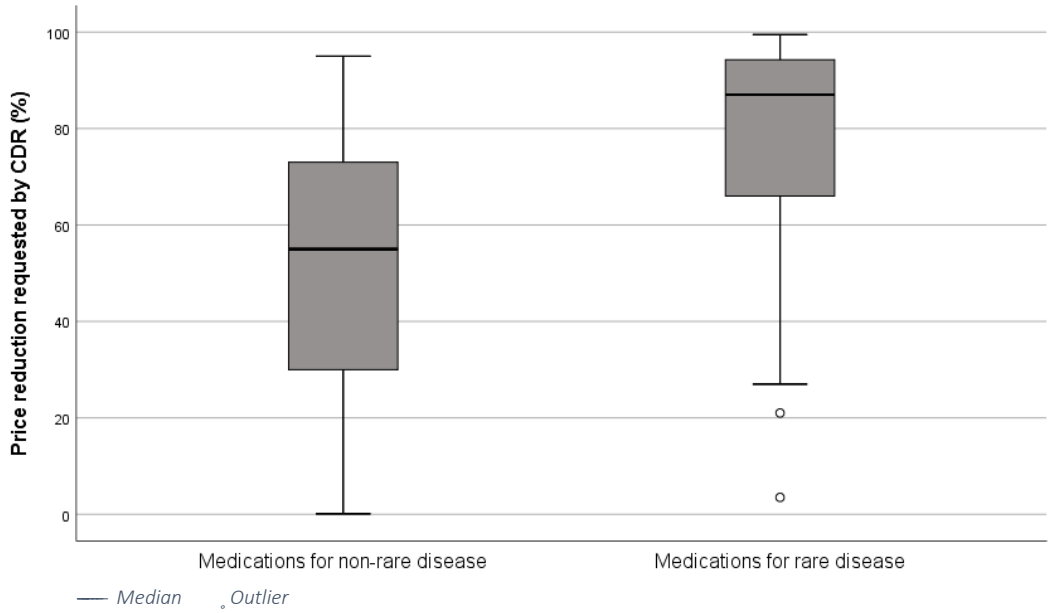

Figure 3. Price reduction requested by CDR based on rare medication and non rare medications

3.6. Willingness-to-pay Threshold mentioned in CDR reports

The CDR reports from 2018 through 2020 indicate that the most requested price reduction was a consequence of the price being greater than the Willingness-to-pay (WTP) threshold of $50,000/QALY. Among the WTP thresholds mentioned in 2018, 2019, and 2020, the most frequent one after $50,000 was $100,000. In 2018 and 2019, other WTP values were also mentioned by CDR, including $25,000, $200,000, $400,000, and $500,000. Over 2021 and 2022, this value threshold was the only WTP threshold mentioned in CDR reports. Figure 4 shows the frequency of WTP thresholds mentioned in CDR reports from 2018 to 2022. It is worth mentioning that in 39 (21.9%) CDR reports analyzed, no specific Willingness-to-pay (WTP) threshold was specified. These

reports were mostly the ones that CDR rejected to assess them as clinically superior and consequently. For these reports, CDR mostly stated that the price of the new intervention should not exceed that of the currently available options.

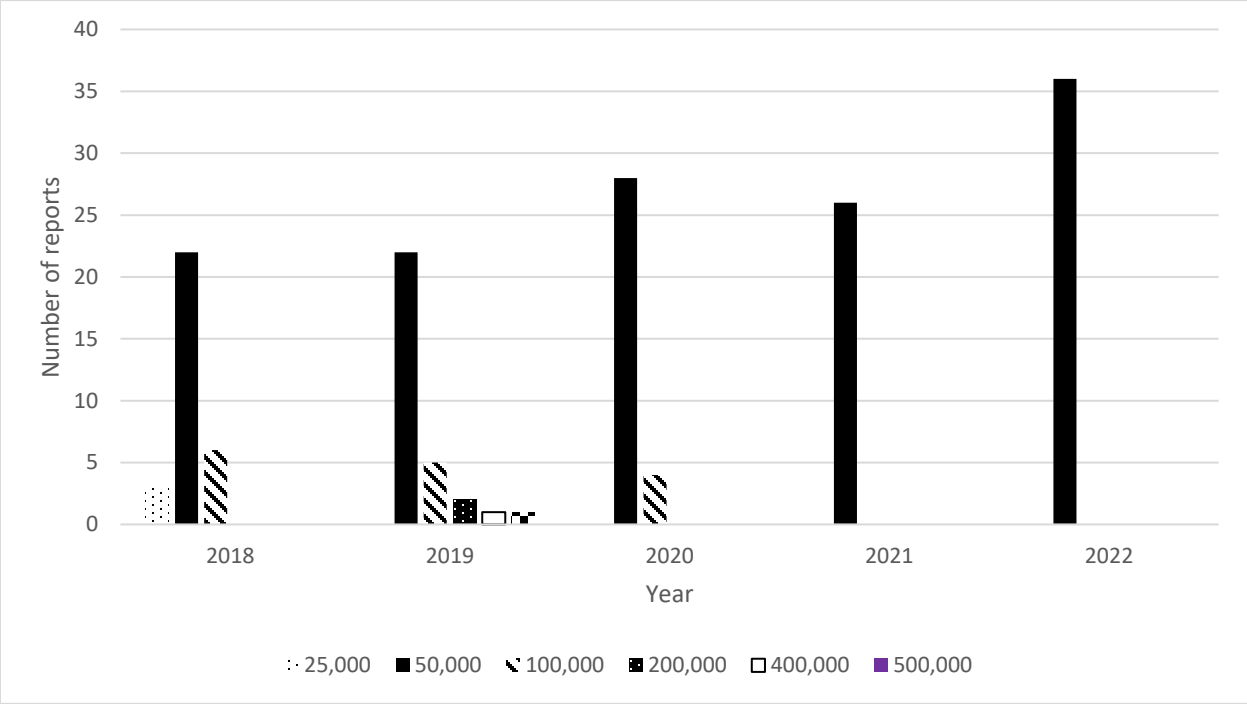

Figure 4. Willingness to pay threshold mentioned in CDR price reduction requests from 2018 to 2022

3.6. CDR recommendations on drugs for rare disease and drugs that address unmet needs

Amongst the reports that claimed to address an unmet need, all 68 reports were recommended with clinical criteria/conditions. For DRD, 47 out of 51 (92%) reports were recommended for reimbursement. In non-DRD, only 100 out of 127 (78.7%) reports received recommendation with clinical criteria/condition. Results of chi-square tests showed an association between addressing an unmet need and DRD with the final CDR recommendation (p value<0.001 and p value<0.05, respectively). The related data is shown in Table 3 and suggests that reports related to DRD have a higher recommendation rate than the one for non-DRD. In addition, it shows that when CDR assessed that a drug had addressed an unmet need, it would have a 100% chance

of receiving "reimburse with clinical criteria/condition". On some occasions, manufacturers claimed to address an unmet need, but CDR refused to accept that. One reason for rejecting these claims would be that the claim is based on a surrogate outcome and not on a more relevant clinical outcome.

Table 3. Final reimbursement recommendations in different therapeutic categories.

	Final reimbursement recommendation								Total reports	
	Reimbursement with clinical criteria/ condition		Reimbursement with clinical criteria		Reimbursement with a price reduction		Do not reimburse			
	N	%	N	%	N	%	N	%	N	%
All submissions	147	82.6	128	71.9	138	77.5	31	17.4	178	100.0
For rare diseases	47	92.2	47	92.2	45	88.2	4	7.8	51	28.7
For non-rare diseases	100	78.7	81	63.8	93	73.2	27	21.3	127	71.3
Addressed an unmet need	68	100.0	65	95.6	68	100.0	0	0.0	68	38.2
For rare disease and addressed an unmet need	28	100.0	28	100.0	28	100.0	0	0.0	28	41.2

"Accepted with clinical criteria/condition" is the explicit phrase that CDR use and it implies any restrictions in price or the clinical use

3.7. Incremental cost and incremental QALY of CDR reanalysis compared to the manufacturer's reports

The median incremental cost provided by the manufacturers was $6,562, with an interquartile interval of -$246 to $127,607, indicating a wide variability in the cost of interventions produced by different manufacturers. The incremental cost in CDR reanalysis was higher, with a median cost of $10,777 and an interquartile interval of $378 to $144,505. The median incremental cost difference was $1,837 with an interquartile interval of -$679 to $25,295. The percentage of incremental cost difference shows that incremental cost increased by 19.6% (interquartile interval -10.1% to 110.5%) in CDR reanalysis compared to the manufacturer's report. Results from the

Wilcoxon rank test indicate that incremental cost is significantly higher in CDR reanalysis than in manufacturer's reports (*p* value<.001).

The median incremental QALY in manufacturers' reports was 0.293 (interquartile interval 0.040 to 1.360) and the median incremental QALY in the CDR reanalysis was 0.111 (interquartile interval 0.011 to 0.765). The difference in incremental QALY values of manufacturers analysis and CDR reanalysis was -0.108 (interquartile interval -0.937 to 0.000) and the percentage of incremental QALY difference value was -50.0%, with an interquartile interval of -83.7% to -1.0%. The Wilcoxon rank test confirms that the median incremental QALY value was significantly lower in the CDR reanalysis compared to the manufacturer's reports (*p* value<.001). The data for incremental costs and QALYs along with the difference of incremental costs and incremental QALYs are demonstrated in Table 4.

3.8. ICURs and differences in ICURs of CDR reanalysis compared to the manufacturer's reports

In the 178 CDR reports, 10 were cost-minimization studies, 13 were cost comparison and 155 were cost utility analyses (CUA). Assessing 67 available ICURs for Manufacturers and CDR's reanalysis, the median ICUR was significantly higher in the CDR reanalysis compared to the manufacturers' reports, $138,658/QALY (interquartile interval 43,203 to 394,076) for manufacturers and $380,251/QALY (interquartile interval 149,197 to 1,347,825) for CDR reanalysis (*p* value<.001). The difference in the median ICURs was $169,299/QALY (interquartile interval 56,040 to 574,073). The difference in ICUR percentage shows that the ICUR values for CDR reanalysis were 150.7% (interquartile interval 32.9% to 352.7%) higher than the manufacturer's

reports. Data related to the incremental costs, incremental QALYs and ICURs are presented in Table 4.

Table 4. Difference in incremental cost, incremental QALY and ICUR of CDR reanalysis compared to the manufacturer's reports

		N	Median	Interquartile interval		Z value	*p* value
				25	75		
Incremental cost							
Manufacturers		167	$6,562	-$246	$127,607	-4.062[b]	*p*<.001
CDR reanalysis		161	$10,777	$ 378	$144,505		
Incremental Cost difference	Value	158	$1,837	-$679	$25,295		
	%	158	19.6%	-10.6%	110.5%		
Incremental QALY							
Manufacturers		147	0.293	0.040	1.360	-6.220[c]	*p*<.001
CDR reanalysis		133	0.111	0.011	0.765		
Incremental QALY difference	Value	133	- 0.108	- 0.937	0.000		
	%	133	-50.0%	-83.7%	-1.0%		
Incremental Cost-utility Ratio							
Manufacturers		67	138,658	43,203	394,076	-6.803[b]	*p*<.001
CDR reanalysis		67	380,251	149,197	1,347,825		
ICUR difference	Value	67	169,299	56,040	574,073		
	%	67	150.7	32.9	352.7		

p value and Z score are produced by the Wilcoxon rank test.

b Based on negative ranks.

c Based on positive ranks.

The distribution of manufacturer's and CDR's ICURs are presented in Figure 5. In 5 reports out of 67 ICUR values, the Manufacturer's ICUR was less than the ICUR values that CDR reanalyzed.

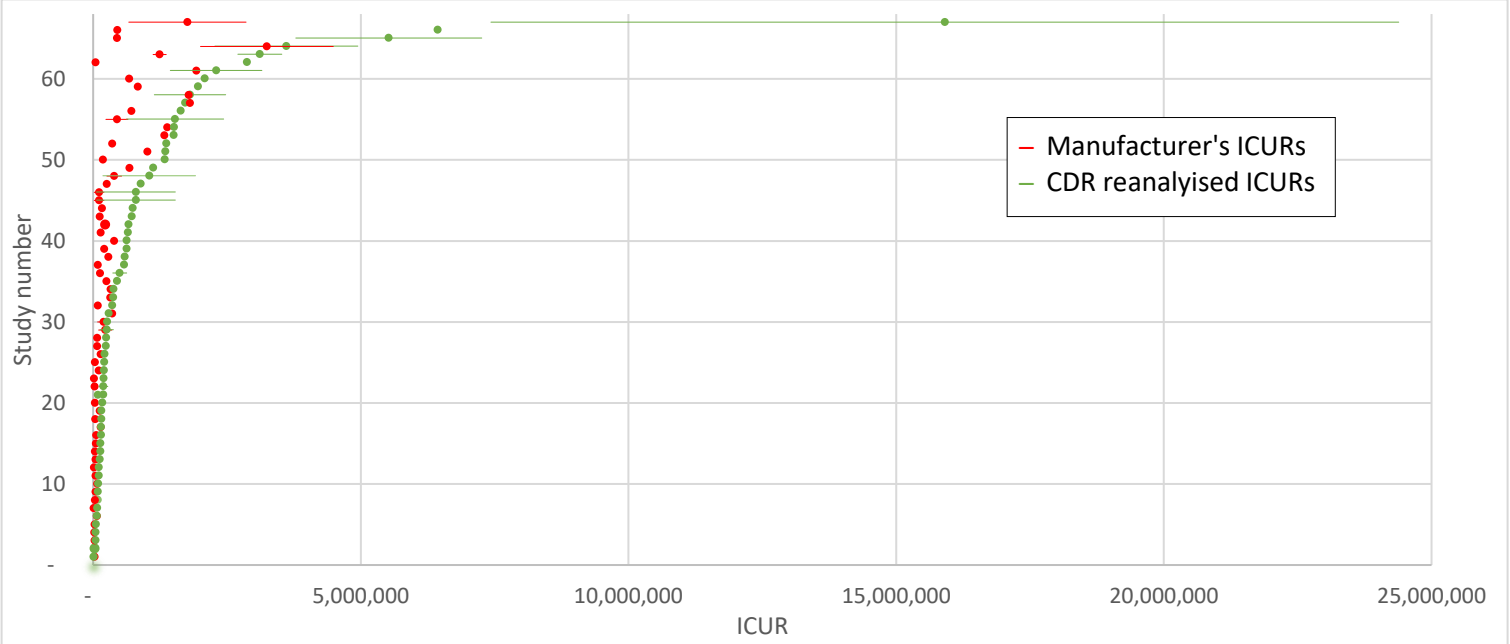

Figure 5. Distribution of ICUR of manufacturers and CDR reanalysis values for 67 CDR reports from 2018 to 2022.

3.9. 3-year Budget impact analysis (BIA) in manufacturer's reports compared to the CDR reanalysis values

Eighty-one studies reported 3-year BIA, among which 44 (54.3%) were assessed in 2022, 30 (37.0%) in 2021, and 7 (8.6%) in 2020.

A 3-year BIA was provided in 81 manufacturers' reports, and CDR reanalyzed 76. The median 3-year BIA for manufacturers was $13,666,621 (interquartile interval 551,393 to 7,264,0799) and significantly lower than the median value of the CDR reanalysis at $19,104,299 (interquartile interval 3,125,747 to 98,198,139) (*p* value<.001). The difference in medians between manufacturers and CDR 3-year BIA was $4,575,102 (interquartile interval of 9,170 to 26,330,512) and the percentage difference was 27.0% (interquartile interval 0.3 to 182). The corresponding values are reported in Table 5.

Table 5. 3-year BIA of manufacturer's report and CDR reanalysis reports along with the difference in values and percentages of 3-year BIAs.

		N	Median	Interquartile interval		Z value	*p* value
				25	75		
Manufacturers 3-year BIA		76	$14,992,393	-$333,901	$77,146,543	-5.043[b]	*p*<.001
CDR reanalysis 3-year BIA		76	$21,515,647	$2,540,979	$99,838,625		
3-year BIA difference	Value	76	$4,575,102	$9,170	$26,330,512		
	%	76	27.0	0.3	182.1		

The distribution of manufacturer's and CDR's 3-year BIAs are demonstrated in Figure 6.

Out of 76 reports, the mean value for 3-year BIA was higher in CDR reanalysis in 66 ones.

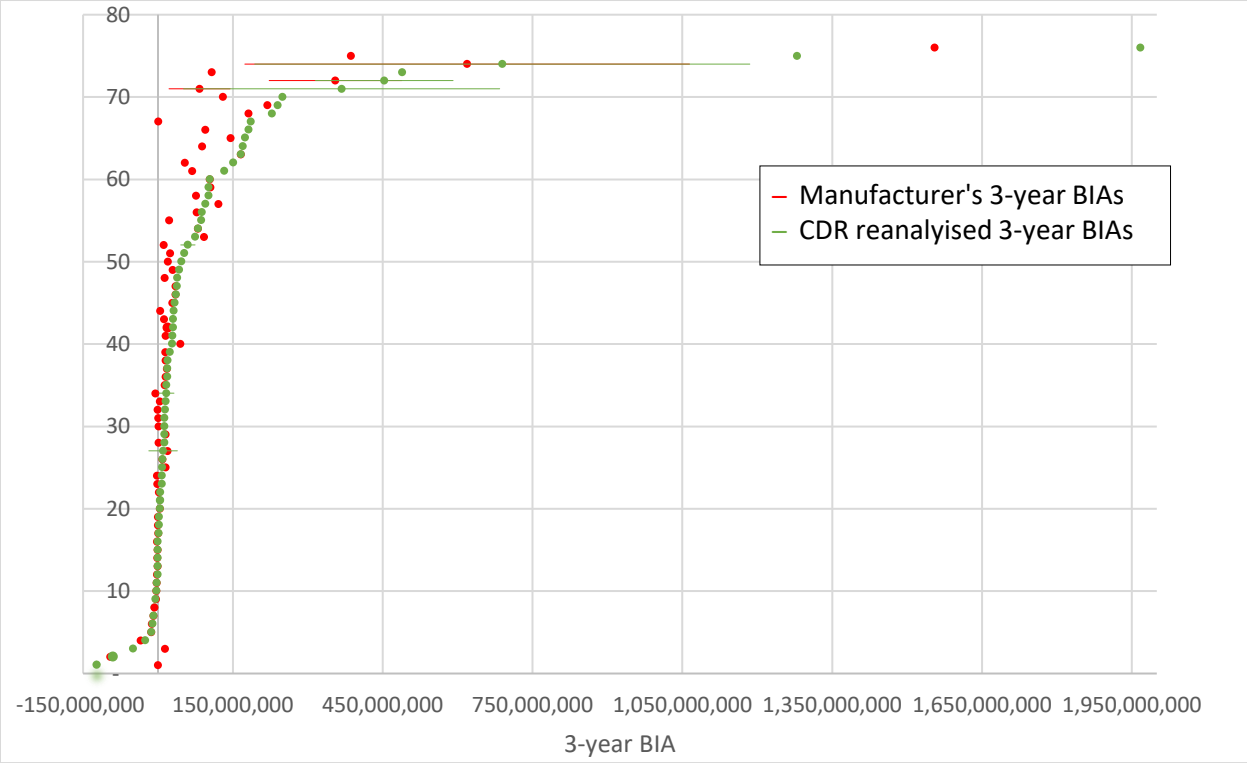

Figure 6. Distribution of 76 3-year budget impact analysis results of manufacturer's reports compared to the CDR reanalysis reports from 2018 to 2022.

3.10. CDR critiques on cost-utility analysis, cost-minimization analysis and cost-comparisons

Almost all the reports (99%) had at least one effectiveness critique. The most common was "critiques on the uncertainty of evidence," which pertained to 96% of the reports, followed by "not enough clinical evidence available (76%), clinical data uncertainty beyond clinical evidence (74%), not a proper population captured (59%).

Eighty-nine percent of studies had at least one structure critique, and the most common structure critique was "model does not reflect current practice," responsible for 59% of all structure critiques.

79% and 69% of the studies had at least one cost and utility score critique, respectively. Complementary data on the critiques, along with their frequencies in different CDR recommendations, are demonstrated in Table S2.

3.11. CDR critiques on uncertainty of evidence

The uncertainty level of evidence is usually specified in the critical appraisal and conclusion part of CDR reports. Among all the reports, 78 (44%) were assessed as highly uncertain, 91 (52%) were assessed as uncertain and only 4 (2%) reports could adequately address the uncertainty levels. Among 31 reports with a "do not reimburse" recommendation, 10 (32%) were rejected merely due to the uncertainty of evidence. However, 64 (44%) of reports with "reimburse with clinical criteria/conditions" recommendation had a high uncertainty level. The chi-square test results show no association between uncertainty level and final CDR recommendation (p value=0.31). Table 6 demonstrates the uncertainty level specified in CDR recommendations based on the final CDR recommendations.

Table 6. Uncertainty level specified in CDR reports based on final recommendation.

General submission characteristics	Final reimbursement recommendation				Total	
	Recommended with clinical criteria/condition		Do not reimburse			
	N	%	N	%	N	%
Highly uncertain	64	44.1	14	45.2	78	44.3
Uncertain	76	52.4	15	48.4	91	51.7
Adequately addressed	2	1.4	2	6.5	4	2.3
Not specified	3	2.1	0	0	3	1.7

"Accepted with clinical criteria/condition" is the explicit phrase that CDR use and it implies any restrictions in price or the clinical use

Although critiques on the uncertainty of evidence were mentioned repeatedly in the CDR critical appraisals, only 7 (4%) of studies mentioned a standardized tool like GRADE for measuring uncertainty of evidence.

3.12. CDR critiques on Risk of Bias

The risk of bias was mentioned in 142 (80%) of the reports but not all domains were discussed. Only 18 (10.1%) mentioned the related tools to measure the risk of bias, with 17 reporting a high risk of bias and 1 with some concerns. In the 4 reports with a "do not reimburse" recommendation, the risk of bias was mentioned as high risk and measured by a standard tool.

3.13. CDR critiques on 3-year BIAs

All the reports had at least one critique on the 3-year BIA analysis. The two most common critiques of 3-year BIA were related to "population of patients" and "percent of market share" responsible, for 80.2% and 73.2% of all 3-year BIA critiques. Complete data on 3-year BIA critiques are shown in Table S2.

3.14. Multivariable regression model

Addressing an unmet need and categorized under DRD entered to the multivariable regression model. Likelihood ratio chi-square test for the logistic regression was 35.59 and the model is significantly predicting the final recommendation significant (p value<.001, Nagelkerke R square= 0.30). Chi-square test results from the Hosmer and Lemeshow test was significant (p value<.001). The results of logistic regression model are demonstrated in table 7. The results show that addressing and unmet need and being in DRD category are predictive variables in receiving "recommendation with criteria/condition" recommendation from CDR.

Table 7. Logistic regression results based on final recommendation.

Variable	Correlation coefficient	P value	Odd ratio	95% CI	
				Lower	Upper
DRD	0.76	0.20	0.14	0.66	0.89
Addressing an unmet need	0.14	0.99	58,391,098	0.00	

4. Chapter four: Discussion

CDR's final recommendations are predominantly influenced by available clinical evidence. Although CDR indicates a price reduction request for expensive medications, no drug was recommended for "do not reimburse" due to its costs alone, and this finding was consistent with previous studies. (7,13,15,18,19) All CDR recommendations were either "reimburse with clinical criteria/conditions" which means price reduction or extra clinical conditions would apply, or "do not reimburse," meaning that no submission was recommended for reimbursement unconditionally by CDR during the period 2018 through to 2022.

CDR reanalysis demonstrated a 20% increase in incremental costs and a 50% decrease in incremental QALY when compared to the manufacturers' report. The median value of ICURs was significantly higher in CDR reanalysis by 51% when compared to manufacturers' analysis. In other words, the differences in both incremental costs and also incremental QALYs between the manufacturers' and CDR's reanalyses resulted in a significant difference in ICURs. In a previous studies similar results were observed. In non-oncology medications, Rocchi et al. surveyed submissions from 2010 to 2017 and reported that the median reanalyzed ICURs for CDR reports was double of mean ICUR submitted by manufacturers. (8) Among three pCORD assessments, similar results were previously show. In the study of Saluja et al., pCODR submissions from 2012 to 2018 were assessed, and it was reported that the median ICURs in manufacturers' reports were lower than the mean pCODR reanalysis ICUR. this study showed that although there was no statistical difference between incremental costs stated by manufacturers and pCODR, it was the lower incremental QALY in the pCODR reanalysis that produced the higher ICURs in the pCODR reanalysis. (10) Raymarkers et al evaluated pCODR reports from 2015 to 2018 and also reported lower ICURs in manufacturers' analysis compared to the pCODR reanalysis, but they found that

the main driver of the ICUR difference was difference in incremental cost.(20) Evaluation of HTA reports for oncology medicines in from 2019 and 2020 showed that pCODR reanalysis demonstrated a 60% reduction in incremental QALY, 183% increase in incremental costs and 82% increase in ICUR compared to manufacturers analysis. (12) In our study, similar to all previously published studies on CADTH reports, manufacturers' ICURs were less than CADTH's reanalyzed ICURs showing that manufacturers have always tended to overestimate the cost-effectiveness of their new interventions.

In this study, almost all the reports had had critiques on effectiveness and more than half of them had critiques on model structure, costing, and utility scores. A high frequency of CDR critiques existed on both accepted and rejected submissions, and the frequency of some of the critiques was even higher in accepted submissions. We have found no specific relationship between the frequency of critiques and final CDR recommendations. One of the most challenging aspects of evaluations in CDR was the uncertainty of evidence which was repeatedly mentioned in the previous literature.(7,15) Although CDR has always tried to enhance transparency and foster evidence-based decision-making, still in less than 4% of the studies GRADE, as a systematic approach to rating the certainty of evidence, was used. The risk of bias, as one of the GRADE domains, was transparently reported using standardized tools in only 10% of CDR reports. However, incorporating all domains of risk of bias in CDR reports could enhance the ability to make more comprehensive comparisons between interventions. We have evaluated the two aforementioned points as weaknesses in CDR's assessments.

Previously, it was reported that pharmaceutical companies might have tendency to under-report the impact of their drugs in BIA. (21) Our analysis also revealed that manufacturers tend to

undermine the budget impact of their new interventions. Although CDR requests price reduction based on the calculated ICURs and it would affect the 3-year BIA of new interventions, we have not seen any rejected submissions due to a high 3-year BIA. This shows that a 3-year BIA is not a critical tool for CDR to judge on final recommendation of new interventions.

From 2018 to 2020, $100,000 was the most frequent WTP threshold mentioned in CDR reports after $50,000. From the beginning of 2021, CDR decided to move on to the more harmonized approach and consequently kept $50,000 as the only WTP threshold for all medications, including DRD. Keeping $50,000 as WTP after 2021 was observed in oncology medication in Canada as well. (22) Some studies stated that addressing an unmet need can override the ICUR threshold in other jurisdictions. However, in this study WTP thresholds were not different between drugs that address an unmet need versus others. (23) Regarding DRD, they had higher general reimbursement recommendations but even from 2018 to 2020, when CDR used higher WTP thresholds in some cases, WTPs used in reports for DRD were not higher than WTP in non-DRD. Although some HTA agencies may consider a higher WTP threshold for DRD, CADTH is strict on its $50,000 WTP threshold. (23)

The median percent price reduction requested for DRD was higher than for non DRD. Since CADTH used to use $100,000 WTP for some ultra rare disease, using $50,000 WTP for all DRD including the one for ultra dare disease resulted in a higher percent price reduction in DRD.(24). One might argue that more than 95% of price reduction requests for 25% of DRD are because these drugs are super-expensive and not affordable for the government. However, even for the 25% lower percentile of Cost per QALY of DRD, up to a 65% price reduction requested from the CDR, which is still significant, and restricting the WTP to $50,000 might have a high impact on

lower-cost DRD. CDR price reduction is only the starting point of the negotiations between the drug plans and the manufacturer and negotiations with pCPA start after CDR's recommendation. pCPA negotiates behind closed doors and publishes the results of its negotiations only as "Concluded without agreement" or "Concluded with an LoA" and does not provide any details about its negotiations.(25) Having mentioned that, some believe that while CDR provides "reimburse with clinical criteria/conditions", the significant percent price reduction might prevent pharmaceutical companies from successfully negotiating with pCPA, specially for DRD. Noting that 1 in 12 Canadians are affected by rare diseases, considering $50,000 WTP for all DRD might jeopardize patients' life by not having access to life-saving medications for rare diseases. (22,26)

4.1. Strengths and Limitations

4.1.1 Strengths

This study possesses several strengths. Firstly, it was a comprehensive review of all 178 reports that CDR published over a 5-year period. Secondly, to minimize errors that could occur with the lengthy nature of each report and the inclusion of both clinical and pharmacoeconomic sections, and the reports were assessed separately by 2 reviewers. Thirdly, it was the first study to assess BIA in CDR reports and to present the BIA values in manufacturers' reports compared to CDR reanalysis reports. As well, there is a discussion of the CDR critiques of the BIA calculated by manufacturers.

4.1.2 Limitations

Our study was limited to the secondary data that is made public by CADTH and we did not have access to the primary data since detailed "confidential" report is done by CDR and provided to manufacturers. We did not review oncology reports and our study is limited to the non-

oncology medications. In addition, considering "reimburse with clinical criteria/condition" as a positive recommendation from CDR was overly optimistic since drugs that receive a recommendation with considerable price reduction might not subsequently receive reimbursement during price negotiation with the pCPA.

4.2. Implications for Health Policymaking in Canada- recommendations for CDR

This study showed that in CDR reports the risk of bias was not reported comprehensively using standardized tools, and there needs to be more metrics for clinical uncertainty in CDR. We suggest CDR report details for the risk of bias evaluation and add a tool (e.g., GRADE) for assessing certainty of evidence.

4.3. Implications for pharmaceutical industry in Canada

In this study, a detailed report containing all critiques on cost-effectiveness models, including critiques for effectiveness, structure, cost, and utility score, were categorized and reported comprehensively. Pharmaceutical companies might use this report to minimize CDR critiques on future submissions. In addition, since most critiques were regarding the clinical evidence, it is recommended that clinical trials designed by pharma industries should prioritize meaningful clinical endpoints to provide more robust data to CDR for clinical and economic evaluations.

5. Chapter five: Conclusion

Higher median incremental costs and lower median incremental QALYs in CDR reanalysis published reports compared to the manufacturers' submitted reports resulted in a higher ICUR in CDR reanalysis. In addition, the 3-year budget impact was higher in CDR analysis, demonstrating that manufacturers tend to underestimate the costs and overestimate the effect of their

medications. WTP mentioned in CDR reports was only $50,000 and the same for all medications in 2021 and 2021.

6. References

1. National health expenditure trends, 2022 — Snapshot [report] [Internet]. Canadian Institute for Health Information. [cited 2023 Mar 21]. Available from: https://www.cihi.ca/en/national-health-expenditure-trends-2022-snapshot

2. Martin D, Miller AP, Quesnel-Vallée A, Caron NR, Vissandjée B, Marchildon GP. Canada's universal health-care system: achieving its potential. Vol. 391, The Lancet. Lancet Publishing Group; 2018. p. 1718–35.

3. Mehrotra A, Dudley RA, Luft HS. What's behind the health expenditure trends? Annu Rev Public Health. 2003;24:385–412.

4. Griffiths EA, Hendrich JK, Stoddart SDR, Walsh SCM. Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. Clin Outcomes Res. 2015;7:463–76.

5. Updates to the Therapeutic Review Framework [Internet]. Canadian Agency for Drugs and Technologies in Health. [cited 2022 Nov 19]. Available from: https://www.cadth.ca/news/updates-therapeutic-review-framework

6. Guidelines for the Economic Evaluation of Health Technologies: Canada. 4th ed. Canada, Ottawa;

7. Rocchi A, Miller E, Hopkins RB, Goeree R. Common Drug Review Recommendations. Pharmacoeconomics. 2012;30(3):229–46.

8. Rocchi A, Mills F. Activities of the pan-canadian pharmaceutical alliance: An observational analysis. J Popul Ther Clin Pharmacol. 2018;25(2):e12–22.

9. Rocchi A, Khoudigian S, Hopkins R, Goeree R. Surrogate outcomes: Experiences at the Common Drug Review. Cost Eff Resour Alloc. 2013;11(1):1–10.

10. Saluja R, Jiao T, Koshy L, Cheung M, Chan KKW. Comparing manufacturer submitted and pan-canadian oncology drug review reanalysed incremental cost-effectiveness ratios for novel oncology drugs. Curr Oncol. 2021;28(1):606–18.

11. Masucci L, Beca J, Sabharwal M, Hoch JS. Methodological Issues in Economic Evaluations Submitted to the Pan-Canadian Oncology Drug Review (pCODR). PharmacoEconomics - Open. 2017;1(4):255–63.

12. Ball G, Levine MAH, Thabane L, Tarride JE. Appraisals by Health Technology Assessment Agencies of Economic Evaluations Submitted as Part of Reimbursement Dossiers for Oncology Treatments: Evidence from Canada, the UK, and Australia. Curr Oncol. 2022;29(10):7624–36.

13. Nagase FNI, Stafinski T, Sun J, Jhangri G, Menon D. Factors associated with positive and negative recommendations for cancer and non-cancer drugs for rare diseases in Canada. Orphanet J Rare Dis. 2019;14(1):1–11.

14. Prescribed drug spending in Canada [product release]. [Internet]. Canadian Institute for Health Information. [cited 2022 Nov 18]. Available from: https://www.cihi.ca/en/prescribed-drug-spending-in-canada

15. Clement FM, Harris A, Jing Jing Li, Yong K, Lee KM, Manns BJ. Using Effectiveness and Cost-effectiveness ro Make Drug Coverage Decisions, A Comparison of Britain, Australia, and Canada. J Am Heart Assoc. 2009;302(13).

16. Budget impact analysis guidelines [Internet]. Patented Medicine Prices Review Board; 2020. Available from: https://www.canada.ca/en/patented-medicine-prices-review/services/reports-studies/budget-impact-analysis-guidelines.html

17. Drugs for Rare Diseases: Evolving Trends in Regulatory and Health Technology Assessment Perspectives. Vol. 2013, Environmental Scan. 2016. p. 0–33.

18. Gotfrit J, Jackson A, Shin JJW, Stewart DJ, Mallick R, Wheatley-Price P. Determinants of the Cancer Drug Funding Process in Canada. Curr Oncol. 2022;29(3):1997–2007.

19. Skedgel C, Wranik D, Hu M. The Relative Importance of Clinical, Economic, Patient Values and Feasibility Criteria in Cancer Drug Reimbursement in Canada: A Revealed Preferences Analysis of Recommendations of the Pan-Canadian Oncology Drug Review 2011–2017. Pharmacoeconomics [Internet]. 2018;36(4):467–75. Available from: https://doi.org/10.1007/s40273-018-0610-0

20. Raymakers AJN, Regier DA, Peacock SJ. Health-related quality of life in oncology drug reimbursement submissions in Canada: A review of submissions to the pan-Canadian Oncology Drug Review. Cancer. 2020;126(1):148–55.

21. Marshall DA, Douglas PR, Drummond MF, Torrance GW, MacLeod S, Manti O, et al. Guidelines for conducting pharmaceutical budget impact analyses for submission to public drug plans in Canada. Pharmacoeconomics. 2008;26(6):477–95.

22. Binder L, Ghadban M, Sit C, Barnard K. Health Technology Assessment Process for Oncology Drugs: Impact of CADTH Changes on Public Payer Reimbursement Recommendations. Curr Oncol. 2022 Mar;29(3):1514–26.

23. Griffiths EA, Hendrich JK, Stoddart SDR, Sean. Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. Clin Outcomes Res. 2015;7:463–76.

24. Rawson NSB. Alignment of health technology assessments and price negotiations for new drugs for rare disorders in Canada: Does it lead to improved patient access? J Popul Ther Clin Pharmacol. 2020;27(February).

25. Husereau D, Dempster W, Blanchard A, Chambers J. Evolution of Drug Reimbursement in Canada : The Pan-Canadian Pharmaceutical Alliance for New Drugs. Value Heal [Internet]. 2023;17(8):888–94. Available from: http://dx.doi.org/10.1016/j.jval.2014.08.2673

26. Canadian Organization for Rare Disorders (CORD). We believe that now is the time to act to provide hope and treatment to Canadians and their families who are impacted by a rare disease. [Internet]. Available from: https://www.raredisorders.ca/our-work/

7. Supplementary appendix

Table S1: Rejection reason mentioned by CDR for 32 studies between 2018 to 2022.

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
Ozurdex	dexamethasone intravitreal implant	Diabetic macular edema	13-Dec-22	09 Diseases of the visual system	Evidence from clinical studies was insufficient to demonstrate a benefit with Ozurdex treatment over other available treatment options in the relevant patient population. It was highly uncertain how Ozurdex compared with anti-VEGF therapy in terms of improving clearness or sharpness of vision due to many limitations in the studies, and there was no evidence comparing Ozurdex with triamcinolone acetonide. Patients identified a need for new treatments that are less invasive and/or require fewer injections, but Ozurdex is administered the same way as other drugs for DME and there was not enough strong clinical evidence to show that patients would experience a tangible benefit from reduced treatment frequency with Ozurdex.
Wakix	pitolisant hydrochloride	Narcolepsy	06-Dec-22	07 Sleep-wake disorders	It is still not known whether Wakix offers any therapeutic benefit over other treatments used for EDS or cataplexy. Two double-blind, randomized, controlled, comparative studies failed to show that Wakix was at least as good as modafinil for treating EDS and did not show any clinical benefit of Wakix over modafinil. One study focused on cataplexy, but this study was versus placebo and results were inconsistent for cataplexy outcomes across the studies. Based on the evidence, the expert committee could not conclude that Wakix met any of the needs that were identified as important to patients, including being more effective and reliable for controlling narcolepsy symptoms and to which patients would be less likely to develop tolerance
Vraylar	cariprazine	Bipolar Disorder	09-Nov-22	08 Diseases of the nervous system	Despite results from 6 clinical trials that showed treatment with Vraylar may improve manic and depressive symptoms associated with bipolar mania and bipolar depression, it is unclear if patients treated with Vraylar in clinical practice would experience the same magnitude of improvement, as patients in the cariprazine studies may not represent the population of patients who will use cariprazine in Canada. Although the results for the 1.5 mg dose of Vraylar suggested an improvement in depressive symptoms, the results for the 3 mg dose of Vraylar were inconsistent across studies; therefore, the CADTH Canadian Drug Expert Committee (CDEC) was not confident that Vraylar would fill a treatment gap. The potential benefit of Vraylar compared to other treatments for bipolar I disorder are unknown. There were no studies directly comparing Vraylar with any other treatments, and the indirect comparative evidence reviewed had many limitations.
Wegovy	semaglutide	Weight management	16-Sep-22	05 Endocrine, nutritional or metabolic diseases	Even though results from 4 clinical trials showed that patients treated with Wegovy for 68 weeks lost more body weight compared to those who received placebo, there was no evidence to show this weight loss translates to improvements in weight-related comorbidities (e.g., cardiovascular complications, osteoarthritis [the most common form of arthritis], and sleep apnea) because they were not studied. Although results showed improvements in health-related quality of life (HRQoL), the minimally important difference (MID) was not met, and it remains unknown if the differences were clinically meaningful. Wegovy is effective for weight loss for up to 2 years with an acceptable side effect profile, but it is unclear whether it meets patient needs for reduced weight-related comorbidities and improved HRQoL.
Spinraza	nusinersen	Spinal Muscular Atrophy	11-Aug-22	08 Diseases of the nervous system	No randomized clinical trials evaluating the efficacy or safety of Spinraza in treatment-naïve adult patients with type II or type III SMA have been conducted. Evidence from 4 observational studies generally suggested that treatment with Spinraza may improve or maintain physical abilities; however, due to the limitations of these studies, it was not possible to conclude that the improvement or maintenance in physical function including movement or strength were a result of Spinraza. There is a need for treatments in adult patients with type II and type III SMA that stabilize disease progression, including the avoidance of

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					using machines to help breathing, improve strength in the upper limbs, and improve health-related quality of life (HRQoL). However, the evidence reviewed did not show that Spinraza would meet any of these needs.
Vraylar	cariprazine	Schizophrenia	10-Aug-22	06 Mental, behavioural or neurodevelopmental disorders	Based on evidence from 5 clinical trials, treatment with Vraylar improved symptoms of schizophrenia or delayed relapse compared with placebo. Vraylar also improved negative symptoms of schizophrenia compared with risperidone. Although these results were statistically significant, it is not clear whether any of these effects result in a clinically meaningful improvement in patient-identified needs. It is not clear whether cariprazine offers any clinical benefits over other treatments that are available for schizophrenia because there were no clinical trials in patients with acute schizophrenia that compared Vraylar with any other treatments. The committee did not have confidence in the results of the indirect comparative evidence because it had too many limitations. There was not enough robust evidence to show that Vraylar filled a treatment gap.
Tavalisse	fostamatinib	Chronic immune thrombocytopenia	05-Apr-22	03 Diseases of the blood or blood-forming organs	Almost no difference with placebo in effectiveness. Treatment with Tavalisse did not significantly reduce bleeding occurrence or severity compared to placebo, and whether treatment with Tavalisse improves health-related quality of life (HRQoL) is not known.
Myinfla	colchicine	Atherothrombotic events in coronary artery disease	08-Mar-22	11 Diseases of the circulatory system	Evidence from 4 clinical trials in patients with coronary artery disease showed that adding Myinfla to standard preventive treatments lowered patients' chances of having major cardiovascular events. There was not enough evidence to show that Myinfla reduced mortality, heart attack and/or stroke, or improved health-related quality of life.
Adtralza	tralokinumab	atopic dermatitis	07-Mar-22	14 Diseases of the skin	Evidence from 3 clinical trials showed that after 16 weeks of treatment, Adtralza was only modestly effective in reducing AD symptoms, including eliminating (or almost eliminating) skin lesions, alleviating itchy skin, and improving quality of life. These modest effects were shown when Adtralza was used alone or in combination with a topical corticosteroid. In another clinical trial in patients with severe AD, Adtralza in combination with topical corticosteroids effectively improved the Eczema Area and Severity Index (EASI) score (a tool used to measure the extent and severity of disease), but this effect was modest. In this same study, treatment with Adtralza in combination with topical corticosteroids did not significantly improve itchy skin than placebo in combination with topical corticosteroids. Results from indirect evidence are inconsistent: 1 indirect comparison suggested that Adtralza is less effective than dupilumab, while the indirect evidence submitted by the sponsor. There is a need for more treatment options for patients whose AD is not controlled despite the use of existing treatments; however, the evidence reviewed did not show that Adtralza would meet this need.
Leqvio	inclisiran	Primary hypercholesterolemia	07-Feb-22	05 Endocrine, nutritional or metabolic diseases	Evidence from 3 clinical trials showed that treatment with Leqvio lowered LDL-C in adults with HeFH or nFH with ASCVD who were already being treated with thehighest possible dose of statins and in those who cannot tolerate treatment with statins. Patients identified a need for treatments that can reduce LDL-C and cardiovascular morbidity and death; however, there was not enough evidence to show that Leqvio would reduce cardiovascular morbidity and death.
Saxenda	liraglutide	Chronic weight management in adults	01-Sep-21	05 Endocrine, nutritional or metabolic diseases	Evidence from 3 studies demonstrated that Saxenda was associated with statistically significant reductions in the body weight compared with placebo after 56 weeks of treatment. No conclusions could be drawn about long-term benefits, particularly for clinically meaningful improvements in comorbidities identifies as priorities by patients such as diabetes, sleep apnea, osteoarthritis and cardiovascular complications. Patients identified a need for treatment that can improve potential obesity-related comorbidities such as diabetes, sleep apnea, and cardiovascular comorbidities. It is not clear whether Saxenda meets these needs.
Zeposia	ozanimod	Multiple Sclerosis, relapsing - remitting	23-Jun-21	08 Diseases of the nervous system	Two randomized, double-blind, active comparator-controlled trials (RADIANCE Part B and SUNBEAM) demonstrated that ozanimod 1 mg was superior to interferon beta1a in terms of reducing the annualized relapse rate (ARR). This was based on a relative ARR reduction of 38% (95% confidence interval [CI], 23% to 49%; $P < 0.0001$) in the RADIANCE Part B study and a relative ARR reduction of 48% (95% CI, 34% to 60%; $P < 0.0001$) in the SUNBEAM study. However, the pooled analyses of the RADIANCE Part B and SUNBEAM studies for time to onset of disability progression (defined as a sustained worsening in Expanded Disability Status Scale [EDSS] of at least 1-point increase), confirmed after 3

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					months and after 6 months, did not demonstrate that there was a statistically significant difference between ozanimod 1 mg and interferon beta1a. The hazard ratios (HRs) for between-group differences for time to onset of disability progression confirmed after 3 months and 6 months were 0.95 (95% CI, 0.68 to 1.33; P = 0.765) and 1.41 (95% CI, 0.92 to 2.17; P = 0.112), respectively. Although patients in the ozanimod 1 mg groups exhibited fewer gadolinium (Gd)-enhanced brain lesions on magnetic resonance imaging (MRI) and fewer new or enlarged hyperintense T2 brain lesions on MRI per year relative to baseline than patients in the interferon beta1a groups in both trials, there was a substantial amount of missing data for these outcomes (24% and 14% of the data were missing in the RADIANCE Part B and SUNBEAM studies, respectively), which increases the uncertainty around the clinical benefits of ozanimod compared with interferon beta1a. There is insufficient evidence to determine if ozanimod offers any meaningful clinical benefits compared with other diseasemodifying treatments (DMTs) for RRMS. Direct comparative evidence for ozanimod 1 mg with DMTs other than interferon beta1a was not identified; however, interferon beta-1a is no longer a routinely used treatment option in current clinical practice, in part because of its modest efficacy, which limits the scope of the results obtained from RADIANCE Part B and SUNBEAM studies. Furthermore, limitations associated with the indirect comparison provided by the sponsor and reviewed by CADTH precluded any conclusions regarding the comparative efficacy and safety advantages of ozanimod with other disease-modifying options for RRMS due to the significant heterogeneity (varying study phase, blinding, diagnostic criteria, publication date, and mean duration of disease) across the included clinical trials. Given the uncertainty in the clinical effectiveness of ozanimod relative to other DMTs, the cost-effectiveness of ozanimod is highly uncertain. CDEC noted that this was highlighted in the CADTH reanalyses, where, in some circumstances, there is no price at which ozanimod would be cost-effective at a $50,000 per QALY threshold.
Corzyna	ranolazine	Stable angina pectoris, adults	27-May-21	11 Diseases of the circulatory system	CDEC reviewed 3 key randomized controlled trials (RCTs: ERICA, CARISA, and TERISA) of ranolazine 1,000 mg twice daily in patients with coronary artery disease and stable angina pectoris. In the ERICA study, the average number of angina episodes per week was reduced to 2.9 (standard error [SE] 0.19) events per week in the ranolazine group, compared with 3.3 (SE 0.22) events per week in the placebo group (P = 0.028). In the TERISA study, the least squares (LS) mean weekly number of angina episodes was reduced to 3.8 [95% confidence interval (CI), 3.6 to 4.1] and 4.3 (95% CI, 4.0 to 4.5) episodes per week for the ranolazine and placebo groups, respectively (P = 0.008). In the CARISA study, ranolazine improved exercise duration on a modified Bruce protocol exercise test relative to placebo with a LS mean difference of 24.0 seconds (SE 11.0, P = 0.03) for the change from baseline measured at trough drug levels. Although results of these trials suggested that ranolazine 1,000 mg twice daily as add-on to standard antianginal drugs reduced angina frequency or improved exercise duration relative to placebo plus standard treatments, the magnitude of benefit on these outcomes was of unclear clinical significance. Further, the key studies were associated with significant limitations, which contribute to the uncertainty in their results. For ERICA and CARISA, uncertainty is also due to significant gaps in the reporting of study methodology, statistical analysis plan, patient characteristics, disposition, and results. The generalizability of the TERISA study is uncertain as the study enrolled an enriched population that were demonstrated to be adherent to the study drug and outcome reporting. CDEC noted that recurrent and sustained angina symptoms would be expected to have an impact on a patient's health-related quality of life (HRQoL) and that an improvement in HRQoL is an important outcome of treatment response in Canadian clinical practice. HRQoL was assessed as a secondary outcome in 2 key studies (ERICA and TERISA). No differences were found between ranolazine and placebo on the disease perception/quality of life domain of the Seattle Angina Questionnaire (SAQ) in the ERICA study, and although there were statistical differences detected in the angina frequency domain, CDEC was unable to draw any conclusions about the clinical relevance of this outcome as there is uncertainty regarding the accepted minimum important difference (MID). No statistically significant differences were found between groups on the change from baseline in the Short Form (36) Health Survey (SF-36) physical component score and mental component score in the TERISA study. Overall, based on the evidence reviewed, the potential benefit of ranolazine on HRQoL remains uncertain. The pharmacoeconomic model submitted by the sponsor was associated with substantial limitations including a lack of sufficient comparative clinical evidence, the inability to reflect disease severity, the relationship between angina frequency and health state utility, and insufficient data to inform treatment response rates. CADTH was unable to address these important limitations. Hence, the cost-effectiveness of ranolazine for the treatment of stable angina remains highly uncertain. CADTH was unable to provide an estimate of the cost-effectiveness of ranolazine for this indication.

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
Spravato	esketamine hydrochloride	Major depressive disorder (MDD), adults	16-Dec-20	06 Mental, behavioural or neurodevelopmental disorders	The results of two four-week randomized controlled trials (RCTs) (TRD3001 and TRD3002) and one randomized withdrawal study (TRD3003) did not demonstrate a consistent statistically significant benefit with esketamine in combination with a newly initiated oral antidepressant compared with placebo in combination with a newly initiated oral antidepressant, in the indicated population. There are several important limitations in the RCTs reviewed that cumulatively result in a high degree of uncertainty regarding the magnitude of the treatment effect of esketamine. Only one of the two four-week induction RCTs demonstrated a statistically significant difference between the esketamine and placebo groups for the primary outcome; there were no direct comparisons of esketamine with other known effective antidepressant therapies; patients in the trials initiated a new oral antidepressant simultaneously with esketamine; unresolved bias remained regarding the potential for unblinding in the trials; the RCTs were of a short duration relative to the duration of MDD; and an enriched population was enrolled across the studies. Due to these limitations, CDEC considered the magnitude of benefit of esketamine in clinical practice to be uncertain and potentially lower than the treatment effects estimated in the three RCTs. The available RCTs were not designed and statistically powered to evaluate several important patient-valued outcomes such as improvements in patients' health-related quality of life, improvements in daily activities or functioning, reduced suicidality, and hospitalizations or emergency department visits. It is unclear if the results of these trials are generalizable to a Canadian population because the patient population enrolled in the trials did not reflect those in whom esketamine would most likely be used in Canadian clinical practice. Esketamine would typically be used in clinical practice later in the treatment pathway than after non-response to only two oral antidepressants, though most participants in the included trials had an inadequate treatment response to two prior oral antidepressants. The overall and long-term balance between the potential benefits and harms of treatment with esketamine are highly uncertain because of limitations with the clinical evidence. A greater proportion of patients who
Cablivi	caplacizumab	Acquired thrombotic thrombocytopenic purpura (aTTP)	26-Aug-20	03 Diseases of the blood or blood-forming organs	Although one phase III, double-blind RCT in adults with acquiredthrombotic thrombocytopenic purpura (aTTP) receiving plasma exchange (PEX) and immunosuppression demonstrated that caplacizumab statistically significantly reduced the time to normalization of platelet count, the study was not designed to assess the effects of caplacizumab on the clinically important outcomes of survival, reduction in organ damage, health care use, or longterm recurrence of aTTP. correlation between time to normalization of platelet count with the aforementioned clinical outcomes. Limitations in the design of the reviewed studies precluded CDEC from determining whether caplacizumab provides clinically meaningful value compared with PEX plus immunosuppression alone. 2 RCTs provided data on the effects of caplacizumab versus placebo for up to two aTTP episodes only. As such, CDEC could not determine caplacizumab's benefit, if any, beyond the duration of the trials. The variability in the natural history of aTTP and the limitations in the design and analysis of RCTs prevented CDEC from identifying a subpopulation of patients with aTTP that is most likely to benefit from treatment with caplacizumab.
Cuvposa	glycopyrrolate	chronic severe drooling, neurologic (pediatric)	24-Jun-20	13 Diseases of the digestive system	Although the results of this RCT suggested that treatement with glycopyrrolate reduced the degree of drooling in some patients compared to treatment with placebo, it is unclear whether the changes observed in drooling represent a clinically meaningful improvement. Specifically, the mTDS is a subjective, unvalidated scale and no information pertaining to the minimal clinically important difference (MCID) for the mTDS was identified in the literature. CDEC determined that it was unlikely that the results of the related RCT is generalizable to patients routinely seen in clinical practice in Canada. The patients included in the study may not be reflective of children with chronic, severe drooling in Canada. There are no comparative studies of glycopyrrolate versus other treatment options used in Canada to reduce severe drooling. Therefore, the comparative effectiveness and safety of glycopyrrolate compared with other treatment options is unknown. A high proportion of adverse effects were observed in the RCT. All 20 patients treated with glycopyrrolate experienced adverse events (AEs), of whom eight patients (40%) had mild AEs, seven (35%) had moderately severe events, and five (25%) had severe AEs. The study was only eight weeks long, which is insufficient to determine the long-term efficacy and potential AEs of treatment for a chronic condition such as severe drooling. Therefore, the long-term safety and efficacy of glycopyrrolate are unknown.

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
Contrave	naltrexone hydrochloride and bupropion hydrochloride	Chronic weight management in adults	27-May-20	05 Endocrine, nutritional or metabolic diseases	In RCTs in adults who were obese or overweight with at least one weight-related comorbidity, those allocated to receive NB experienced an average of 3% to 5% reduction in body weight after 28 or 56 weeks compared with placebo. The relevance of a 3-5% weight loss over a short time period on obesity, a chronic long-term condition, and on clinical conditions such as diabetes, hypertension, other cardiovascular disease or other conditions such as sleep apnea remains uncertain. None of the trials demonstrated resulted in clinically meaningful improvements in weightrelated comorbidities, symptoms of importance to patients (fatigue, pain, and impaired productivity, sleep, and mobility), or health-related quality of life. Sustained weight loss over many years is necessary to influence many chronic disease outcomes. Given the short duration of the included trials and the early termination of a long-term cardiovascular safety study, the long-term benefits and risks of NB are unknown.
Benlysta	Belimumab	systemic lupus erythematosus	22-Apr-20	04 Diseases of the immune system	In one 52-week double-blind, randomized controlled trial (RCT) that compared belimumab SC treatment to placebo in patients with active SLE on standard background therapy (the BLISS-SC study), a statistically significant greater proportion of belimumab SC-treated patients achieved a response based on the SLE responder index (SRI, 61% versus 48%; odds ratio [OR] 1.68, 95% confidence interval [CI], 1.25 to 2.25) and a smaller proportion of belimumab SC-treated patients experienced a severe flare (11% versus 18%; hazard ratio 0.51, 95% CI, 0.35 to 0.74). Despite being statistically significant, the improvement in the response rate (13% higher in belimumab SC-treated patients versus placebo-treated patients) was considered by CDEC to be relatively modest. In the BLISS-SC study, belimumab SC treatment did not statistically significantly reduce the proportion of patients who were able to reduce the dose of prednisone used to 7.5 mg per day or less. Furthermore, the BLISS-SC study failed to assess the effect of belimumab SC on several other outcome measures considered to be important to patients, including health-related quality of life (HRQoL) and activities of daily living. Patients enrolled in the BLISS-SC study were heterogenous with respect to their standard therapies received at time of enrolment. Furthermore, it is unclear to what extent these treatments were optimized. Therefore, CDEC was unable to identify a subpopulation of patients with SLE who might be more likely to respond to belimumab SC.
Onstryv	safinamide	Parkinson's disease	25-Mar-20	08 Diseases of the nervous system	In two RCTs safinamide did not show a clinically meaningful improvement over placebo in change from baseline to Week 24 in ON time, the primary outcome in each trial. For the assessment of daily OFF time at Week 24 in SETTLE, the difference in change from baseline between safinamide 50 mg/day to 100 mg/day and placebo was –1.03 hours in favour of safinamide. In Study 016, the difference in change from baseline between safinamide 50 mg/day and placebo was –0.6 hours in favour of safinamide. The difference in change from baseline between safinamide 100 mg/day and placebo in Study 016 was –0.6 hours in favour of safinamide. The difference in SETTLE met the published minimally important difference (MID) in OFF time (-1 to -1.3 hours), however, the upper bound of the 95% confidence interval is outside the MID range. The relative efficacy of safinamide compared with other add-on treatments used for PD is unclear due to major limitations associated with the manufacturer-submitted indirect treatment comparison (ITC) and one published ITC by Binde et al., 2018. There is no evidence that safinamide addresses an unmet need that is not already addressed by other add-on treatments currently reimbursed for the treatment of PD, including better management of OFF episodes, improved quality of life, or improved non-motor outcomes relevant to patients such as sleep, pain, mood and constipation.
Iluvien	fluocinolone acetonide intravitreal implant	diabetic macular edema	26-Sep-19	09 Diseases of the visual system	There was no direct evidence comparing fluocinolone acetonide intravitreal implant with other active treatments used in Canada for the treatment of adult patients with DME. The two phase III randomized controlled trials (RCTs) identified in the systematic review were designed to compare fluocinolone acetonide with sham treatment. Inconsistent results in measures of best-corrected visual acuity (BCVA) were observed within each FAME study and across the trials. No between-groups difference in the proportion of patients with a worsening from baseline of three or more steps in the Early Treatment Diabetic Retinopathy Study (ETDRS) multi-step eye scale of diabetic retinopathy was observed in either trial. Results of secondary outcomes were either not statistically significant or not consistent between trials. In both trials, a higher percentage of patients in the fluocinolone acetonide 0.2 mcg/day group experienced ocular-related adverse events than in the sham group. Increased intraocular pressure was reported in a greater percentage of patients treated with fluocinolone acetonide than in sham-treated patients . There are insufficient data to assess the safety and efficacy of fluocinolone acetonide in patients who would use fluocinolone acetonide

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					intravitreal implant as a second-line therapy. Anti-VEGF therapy is considered first-line treatment for DME in Canada, but responses of patients previously treated with anti-VEGFs to fluocinolone acetonide is unknown. The number of patients with prior exposure to anti-VEGF treatment (between 4.2% and 7.5%) was lower in the FAME trials than in Canadian clinical practice. In addition, patients enrolled in the FAME trials were not required to have been previously treated with a course of corticosteroids, as stipulated in the Health Canada indication.
Xermelo	telotristat	carcinoid syndrome	26-Jun-19	05 Endocrine, nutritional or metabolic diseases	In one double-blind, placebo-controlled RCT, telotristat was associated with a statistically significant greater reduction in the number of daily bowel movements compared to placebo over 12 weeks, but the magnitude of this difference between telotristat and placebo was of uncertain clinical relevance. Telotristat failed to improve symptoms of importance to patients that are associated with diarrhea such as urgency and did not improve a variety of health-related quality of life subscales associated with this condition (fatigue, body image, pain, impact on finances, and social/cognitive functioning). Moreover, telotristat did not improve other symptoms associated with carcinoid syndrome, namely abdominal pain and flushing. Patients in the RCT received concomitant SSA therapy (octreotide and lanreotide). Some patients had their SSA therapy optimized with escalated doses while others continued to receive standard doses. It is unclear what effects telotristat would have in patients with carcinoid syndrome diarrhea that is inadequately controlled by optimized SSA therapy.
Jublia	efinaconazole	Onychomycosis	23-May-19	14 Diseases of the skin	Several medications used to treat onychomycosis in Canada are reimbursed by the Common Drug Review (CDR)-participating drug plans, including terbinafine and itraconazole, and there is no evidence to suggest that topical efinaconazole fulfills an unmet need in treating this condition. There are no randomized controlled trials (RCTs) in which efinaconazole has been compared directly with other treatments used to treat onychomycosis in Canada, and the results of an indirect treatment comparison (ITC), while having limitations, suggested that topical efinaconazole was less effective than oral terbinafine 250 mg daily and itraconazole 200 mg daily at inducing mycologic cure of onychomycosis.
Eucrisa	crisaborole	atopic dermatitis	27-Mar-19	14 Diseases of the skin	Two RCTs of patients demonstrated that treatment with crisaborole was significantly more likely than the vehicle to achieve an Investigator Static Global Assessment (ISGA scoring) score of 0 to 1 (clear or almost clear) with at least a 2-grade improvement from baseline at Day 29. However, the benefits shown in both RCTs were not compared with standard treatments such as topical corticosteroids (TCS), topical calcineurin inhibitors (TCIs), systemic immunomodulating drugs, or phototherapy. Therefore, there is no direct evidence demonstrating comparative efficacy for crisaborole versus other standard treatments for AD. Both RCTs were not representative of the populations for whom the reimbursement request was made: patients two years and older with mild-to-moderate AD who have failed or are intolerant to a TCS. No evidence related to the reimbursement request population was available in either RCT. The results of both manufacturer-submitted and published network meta-analyses (NMAs) show there were no statistically significant differences found between crisaborole and TCIs (pimecrolimus or tacrolimus) for the proportion of patients achieving ISGA score of 0 to 1 (clear or almost clear). Limitations to the NMAs included the limited number of trials available to inform the network and that only comparisons versus TCIs (pimecrolimus and tacrolimus) were reported. Subgroup analyses based on age were not conducted, the reporting of only one efficacy outcome (achieving an ISGA score of 0 to 1) to assess comparative treatment effects was used in the analysis, and no quantitative assessment of comparative safety was done.
Segluromet	ertugliflozin and metformin hydrochloride	Diabetes mellitus, Type 2	23-Jan-19	05 Endocrine, nutritional or metabolic diseases	Although evidence from two RCTs demonstrated that ERT (as add-on combination with metformin, or with metformin plus sitagliptin) statistically significantly improved glycated hemoglobin (A1C) after 26 weeks of treatment compared with placebo, ERT has not been demonstrated to have benefits on longer-term clinical outcomes such as reducing major cardiovascular events, which has been reported for some other sodium-glucose cotransporter-2 (SGLT2) inhibitors. Because there is no evidence of a cardiovascular benefit for ERT, if patients without a high risk of cardiovascular events were to initiate therapy with ERT/MET, those that subsequently experienced an increase in the risk of a cardiovascular event would need to be switched to an alternative treatment for which there is evidence of cardiovascular benefit. CDEC considered such a treatment strategy to be both difficult to implement and to be associated with an increased risk of harm to patients. There are limited data to compare ERT/MET with other anti-hyperglycemic FDCs available for the treatment of adult patients with type 2 diabetes mellitus. One RCT (SU study) suggested that ERT, as add-on to metformin, was noninferior to glimepiride plus metformin for the

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					change from baseline in A1C after 52 weeks, but for the ERT 15 mg once daily dosage form only. In another RCT (FACTORIAL), statistically significant short-term (26 weeks) improvements in A1C, body weight, and systolic blood pressure were observed for ERT plus sitagliptin, as add-on therapy to metformin, versus sitagliptin plus metformin. A manufacturer-submitted indirect treatment comparison of ERT versus other SGLT2 inhibitors and placebo suggested that ERT in combination with metformin for the treatment of type 2 diabetes mellitus is likely more efficacious than placebo; however, concrete conclusions could not be drawn with respect to the comparative efficacy with other SGLT2 inhibitors added onto metformin, or the relative safety of ERT/MET. Therefore, there is uncertainty regarding the long-term comparative benefits and safety of ERT/MET versus other treatments that are available for patients with type 2 diabetes mellitus, and there is no evidence that ERT/MET fulfills an unmet need in the treatment of patients with type 2 diabetes mellitus.
Steglatro	ertugliflozin	Diabetes mellitus, Type 2	23-Jan-19	05 Endocrine, nutritional or metabolic diseases	Although evidence from three double-blind randomized controlled trials (RCTs) demonstrated that ERT (as monotherapy, or as add-on combination with metformin, or with metformin plus sitagliptin) statistically significantly improved glycated hemoglobin (A1C) after 26 weeks of treatment versus placebo, ERT has not been demonstrated to have benefits on longer-term clinical outcomes such as reducing major cardiovascular events, which has been reported for some other sodium-glucose cotransporter-2 (SGLT2) inhibitors. Because there is no evidence of a cardiovascular benefit for ERT, if patients without a high risk of cardiovascular events were to initiate therapy with ERT, those that subsequently experienced an increase in the risk of a cardiovascular event would need to be switched to an alternative treatment for which there is evidence of cardiovascular benefit. CDEC considered such a treatment strategy to be both difficult to implement and to be associated with an increased risk of harm to patients. There are limited data to compare ERT with other antihyperglycemic drugs available for the treatment of adult patients with type 2 diabetes mellitus. One RCT (VERTIS SU) suggested that ERT, as add-on to metformin, was noninferior to glimepiride plus metformin for the change from baseline in A1C after 52 weeks, but only at 15 mg once-daily dosage. In another RCT (VERTIS FACTORIAL), statistically significant short-term (26 week) improvements were reported in A1C, body weight, and systolic blood pressure for ertugliflozin plus sitagliptin as add-on therapy to metformin, versus sitagliptin plus metformin. A manufacturersubmitted indirect treatment comparison of ERT versus other SGLT2 inhibitors and placebo suggested that ERT as monotherapy or in combination with metformin is likely more efficacious than placebo; however, concrete conclusions could not be drawn with respect to the comparative efficacy or relative safety of ERT when compared with other SGLT2 inhibitors. Therefore, there is uncertainty regarding the long-term comparative benefits and safety of ERT versus other treatments that are available for patients with type 2 diabetes mellitus, and there is no evidence that ERT fulfills an unmet need in the treatment of patients with type 2 diabetes mellitus.
Ozanex	ozenoxacin	Impetigo	24-Oct-18	01 Certain infectious or parasitic diseases	No high-quality direct evidence comparing ozenoxacin with topical antibiotics used in Canada to treat impetigo was identified. The two phase III, placebo-controlled randomized controlled trials (RCTs) of patients with impetigo identified in the systematic review (Study P-880 and Study P-881) were designed to compare ozenoxacin with vehicle (placebo). The two manufacturer-submitted indirect treatment comparisons (ITCs), of ozenoxacin versus sodium fusidate and ozenoxacin versus mupirocin, had several limitations, including the availability of only one study per direct comparison in each ITC. Limitations of the ozenoxacin versus sodium fusidate ITC included the use of a post hoc end point in Study P880, and there was a lack of information on the use of concomitant antimicrobial therapies in the other trial. The ozenoxacin verus mupirocin ITC was limited by a high risk of attrition bias and small sample size in one study, and between study heterogeneity in terms of study design and patient characteristics. Consequently, the comparative efficacy and safety of ozenoxacin versus other available topical antibiotic treatments for impetigo remains uncertain. CDEC determined that the reviewed clinical trials do not provide sufficient evidence that ozenoxacin would fulfill an unmet need in the treatment of impetigo.
Ozurdex	dexamethasone	Diabetic macular edema	24-Oct-18	09 Diseases of the visual system	There was no high-quality direct evidence comparing dexamethasone implant with other active treatments used in Canada for the treatment of adult patients with DME who are pseudophakic (e.g. laser therapy, intravitreal steroid, or anti-VEGF therapies) identified. The two phase III sham-controlled randomized controlled trials (RCTs) identified in the systematic review (MEAD-010 and MEAD-011) were designed to compare dexamethasone implant with sham. Compared with sham, the mean change from baseline in best corrected visual acuity (BCVA) in the pre-specified subgroup of patients with

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					DME who were pseudophakic did not exceed a 10-letter improvement (5.9 letters and 3.6 letters in MEAD-010 and MEAD-011 respectively), and between-treatment differences in the proportion of patients achieving a ≥ 15-letter improvement, favouring dexamethasone were modest; 18.1% (95% confidence interval [CI], 0.8 to 35.4) and 6.0% (95% CI, –5.7 to 17.8) in MEAD-010 and MEAD-011, respectively. Further, the lack of stratification by lens status and failure to control for multiplicity results in uncertainty regarding the magnitude of benefit. Based on a pooled analysis of the MEAD trials, a higher percentage of pseudophakic patients in the dexamethasone group experienced adverse events, such as elevated intraocular pressure , compared with the sham group (29.4% and 9.0% ,respectively), which is consistent with the adverse event profile of intravitreal steroid therapies. The manufacturer-submitted indirect comparison (IDC) was limited by clinical and methodologic heterogeneity, and phase II studies comparing dexamethasone implant with ranibizumab or bevacizumab were not stratified by lens status and subgroup analyses for the pseudophakic population were not controlled for multiplicity and lacked statistical power. Thus, no conclusions could be made regarding the comparative efficacy and safety of dexamethasone implant versus relevant comparators for the treatment of adult patients with DME who are pseudophakic. Between 9.1% and 16.0% of patients in the subgroup of pseudophakic patients included in the MEAD-010 and MEAD-011 trials, respectively, had prior experience with anti-VEGF therapy. However, the responses of these patients to anti-VEGF treatment. Therefore, there are insufficient data to assess the safety and efficacy of dexamethasone in patients who would use dexamethasone implants as second-line therapy (e.g., have had an inadequate response to or did not tolerate prior anti-VEGF therapy).
Nucynta	tapentadol hydrochloride	pain, severe	23-Oct-18	21 Symptoms, signs or clinical findings, not elsewhere classified	All eight randomized controlled trials (RCTs), including five in non-cancer pain (comparing tapentadol ER with oxycodone controlled release [CR] or oxycodone plus naloxone prolonged release [PR]) and three in cancer-related pain (comparing tapentadol ER with oxycodone CR, morphine sustained release [SR], or morphine CR), contained numerous limitations that impact the validity of the results. These limitations included high and unbalanced withdrawal rates with corresponding missing data, short duration, lack of blinding of active controls, lack of statistical testing or control of multiplicity, and differential use of concomitant rescue analgesics. These limitations may have biased the comparative measure of effect in reducing pain intensity and resulted in significant uncertainty regarding the magnitude of the effect of tapentadol. Despite evidence from the aforementioned trials to suggest that tapentadol ER may be associated with fewer adverse events (AEs), AEs were reported in more than 65% of all active treatment groups in seven of eight trials. In addition, the observed differences in gastrointestinal AEs that favoured tapentadol ER were potentially impacted by the rapid titration of opioids and the apparent lack of specific bowel management regimens in the trials, neither of which are reflective of clinical practice in the treatment of chronic pain. Direct comparisons were not available for several long-acting opioids, such as oral hydromorphone or methadone, transdermal fentanyl, or transdermal or buccal film buprenorphine, tramadol ER, or codeine CR. Although two indirect treatment comparisons (ITCs) were provided, these were associated with serious limitations and failed to provide clarity regarding the relative benefit of tapentadol ER versus other treatments for chronic pain requiring long-term opioid treatment. There was insufficient evidence to suggest that tapentadol ER fulfills an unmet need within the current treatment landscape for chronic pain in reducing the potential for opioid use disorder, misuse, overdose, or diversion, compared with other available opioids.
Orkambi	lumacaftor/ivacaftor	Cystic Fibrosis, F508del CFTR mutation in patients 6 years and older	26-Sep-18	12 Diseases of the respiratory system	Although two double-blind, randomized controlled trials (RCTs) (TRAFFIC [N = 559] and TRANSPORT [N = 563]) demonstrated that treatment with LUM 400 mg every 12 hours/IVA 250 mg every 12 hours (L400/IVA) was associated with statistically significant absolute improvements in per cent predicted forced expiratory volume in one second (ppFEV1) compared with placebo, the magnitude of improvement (2.6% to 3.0%) was of uncertain clinical significance. In addition, responder analyses demonstrated that the majority of L400/IVA-treated patients (73%) failed to achieve an absolute improvement of at least 5% in ppFEV1. The manufacturer conducted a matched-registry cohort analysis that suggested the slope of decline in lung function was reduced in patients who were treated with L400/IVA in the PROGRESS study compared with a matched cohort of patients from a US registry (–1.33% versus –2.29% per year over a two-year period). Due to limitations in the analysis, concerns regarding the comparability of the patients from the clinical trials and those from the registry, and issues regarding the generalizability of US registry patients with Canadian patients with CF, it is uncertain if treatment with L400/IVA would have a similar impact

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					on the rate of lung function decline in Canadian patients. L400/IVA was associated with a lower rate of pulmonary exacerbations compared with placebo after 24 weeks of treatment in the TRAFFIC and TRANSPORT trials; however, the results could not be considered statistically significant because the hierarchical statistical analysis plan used in both studies failed to demonstrate statistical significance at a higher order comparison. As well, the data for pulmonary exacerbations were limited by the relatively short duration of the trials and the absence of independent adjudication of exacerbation events. The included RCTs failed to consistently demonstrate that treatment with L400/IVA is associated with statistically significant improvements in body mass index (BMI), body weight, and height in patients at least 12 years of age. Although a statistically significant improvement in BMI was reported in TRANSPORT, the magnitude of improvement was of uncertain clinical significance. In patients aged six years to 11 years, treatment with LUM 200 mg every 12 hours/IVA 250 mg every 12 hours (L200/IVA) was not associated with statistically significant improvements in nutritional BMI, BMI-for-age z score, weight, weight-for-age z score, height, or height-for-age z score. In patients aged six years to 11 years of age, L200/IVA was associated with a statistically significant improvement in lung clearance index 2.5% (LCI2.5) compared with placebo after 24 weeks of treatment (absolute reduction of –1.09). The clinical significance of this finding is uncertain as the minimally clinically important difference (MCID) has not been established for this end point, its validity as a surrogate marker for respiratory exacerbations is unknown, and it is not currently used in Canadian clinical practice. Treatment with L200/IVA resulted in an improvement in ppFEV1 after 24 weeks of treatment compared with placebo (2.4%); however, the clinical significance of this result is uncertain. Treatment with L200/IVA was not associated with a statistically significant improvement in the rate of pulmonary exacerbations in patients six years to 11 years of age. There were no statistically significant improvements in health-related quality of life using the Cystic Fibrosis Questionnaire – Revised (CFQ-R) respiratory domain scores with L400/IVA or L200/IVA versus placebo at 24 weeks.
Viberzi	eluxadoline	Irritable bowel syndrome with diarrhea	24-Aug-18	13 Diseases of the digestive system	The results of two phase III, double-blind, randomized, placebo-controlled, parallel-group, trials (IBS-3001, N = 1,281; and IBS-3002, N = 1,146) demonstrated a statistically significant improvement in the primary composite outcome (daily pain response and daily stool consistency response) for eluxadoline (75 mg twice daily and 100 mg twice daily) compared with placebo between baseline and 12 and 26 weeks. However, only approximately one-third of patients in the eluxadoline treatment arm were considered responders for this composite end point, which was driven primarily by stool consistency responders. There were no statistically significant differences between eluxadoline and placebo for the percentage of daily pain responders in either study between baseline and 12 or 26 weeks. Quality of life comparisons using the least squares mean difference in irritable bowel syndrome quality of life (IBS-QoL) questionnaire results showed a statistically significant difference between the active groups and placebo except at week 26 and week 30 in the 100 mg eluxadoline arm in IBS-3002. However, there was no clear benefit of eluxadoline on patients' quality of life based on the IBS-QoL questionnaire when measured as the percentage of patients who met the pre-specified minimal clinically important difference (MCID) threshold of a 14-point difference, except in the IBS-3001 eluxadoline 100 mg group at week 52. The lack of control for multiple statistical testing for outcomes other than the primary composite outcome and the high percentage of patients discontinuing from the studies further limit the ability to interpret the findings. Eluxadoline use was associated with higher rates of withdrawals due to adverse events compared with placebo in both trials. Pancreatitis was reported in seven patients, all of whom were in the eluxadoline treatment groups. No direct or indirect comparative evidence is available to assess the clinical benefit of eluxadoline versus other pharmacological agents commonly used to treat IBS-D.
Dupixent	dupilumab	atopic dermatitis	27-Jun-18	14 Diseases of the skin	No evidence was available that compared dupilumab with other drugs commonly used in the treatment of AD. The four phase III, placebo-controlled, randomized controlled trials (RCTs) (three 16-week trials [SOLO 1, SOLO 2, and LIBERTY AD CAFÉ] and one 52-week trial [LIBERTY AD CHRONOS]) reviewed were not designed to compare dupilumab with other drugs commonly used in the treatment of atopic dermatitis. Although these trials demonstrated that a statistically significantly greater percentage of patients had improvements in AD severity, symptoms, and quality of life with dupilumab treatment compared with placebo, the magnitude of clinical benefit with dupilumab compared with existing alternative treatments is unknown. There are several notable gaps in the clinical evidence regarding dupilumab, including data to assess the long-term safety of dupilumab,

Brand Name	Generic Name	Therapeutic Area	Date Recommendation Issued	ICD-11Therapeutic area	Rejection reason
					concerns with the generalizability of the trial results to patients who would be expected to use dupilumab in clinical practice, and an absence of efficacy and safety data for the use of dupilumab in patients where topical prescription therapies are not advisable.

Table S2: CDR Critiques on cost-utility, cost-minimization, and cost comparison reports of manufacturers in "Recommended with clinical criteria/condition" group, "Do not reimburse" group and in total reports.

	Final reimbursement recommendation				Total	
	Recommended with clinical criteria/condition		Do not reimburse			
	N	%	N	%	N	%
Effectiveness critiques						
Critiques on uncertainty of evidence	140	95.2	31	100.0	171	96.1
Not enough clinical evidence available	112	76.2	24	77.4	136	76.4
Clinical data uncertainty beyond clinical evidence	111	75.5	21	67.7	132	74.2
Not a proper population captured	89	60.5	16	51.6	105	59.0
Limitation in ITC or NMA	65	44.2	20	64.5	85	47.8
Inappropriate comparator was used in the base case analysis	67	45.6	17	54.8	84	47.2
AE not captured correctly	54	36.7	13	41.9	67	37.6
Incorrect or double counting mortality	42	28.6	8	25.8	50	28.1
Evidence with low number of participants	40	27.2	5	16.1	45	25.3
Critiques on using surrogate outcome	23	15.6	9	29.0	32	18.0
Critiques on extrapolation	16	10.9	1	3.2	17	9.6
Inappropriate cut off points	13	8.8	2	6.5	15	8.4
Using interim data instead of final ones	9	6.1	0	0.0	9	5.1
Critiques on risk equation	6	4.1	2	6.5	8	4.5
Other effectiveness critiques	71	48.3	21	67.7	92	51.7
Any effectiveness critiques	146	99.3	31	100.0	177	99.4
Structure critiques						
Model does not reflect current practice	86	58.5	19	61.3	105	59.0
Model structure did not capture some important aspects	59	40.1	14	45.2	73	41.0
Lack of face validity or CADTH could not validated the model	60	40.8	12	38.7	72	40.4

Lack of transparency	39	26.5	14	45.2	53	29.8
Model was not flexible enough for CADTH to change some aspects	41	27.9	9	29.0	50	28.1
Programming error	35	23.8	13	41.9	48	27.0
Model is based on the treatment response instead of health states	24	16.3	2	6.5	26	14.6
Other structure critiques	83	56.5	20	64.5	103	57.9
Any structure critiques	130	88.4	28	90.3	158	88.8
Utility critiques						
Utility score(s) do not reflect Canadian ones	58	39.5	6	19.4	64	36.0
Utility scores not related to health states	19	12.9	2	6.5	21	11.8
Other utility critiques	71	48.3	16	51.6	87	48.9
Any utility critiques	103	70.1	20	64.5	123	69.1
Cost critiques						
Unacceptable extra costs (like hospitalization, administration, or other charges)	59	40.4	6	19.4	65	36.7
Missing or underestimation of some costs	47	32.0	12	38.7	59	33.1
Inappropriate list prices for the comparator(s)	29	19.7	5	16.1	34	19.1
Other cost critiques	55	37.7	13	41.9	68	38.4
Ay cost critiques	116	78.9	24	77.4	140	78.7

Table S3: CDR Critiques on 3-year budget impacts analysis of manufacturer's reports in "Recommended with clinical criteria/condition" group, "Do not reimburse" group and in total reports.

	Final reimbursement recommendation				Total	
	Recommended with clinical criteria/condition		Do not reimburse			
	N	%	N	%	N	%
Population of patients	58	84.1	7	58.3	65	80.2
percent of market share	51	73.9	9	69.2	60	73.2
Critiques due to the clinical data uncertainty	25	36.2	4	30.8	29	35.4
Percent of patients covered by public plan	23	33.3	4	30.8	27	32.9
Intervention cost/intervention related costs	22	31.9	1	7.7	23	28.0
Cost of comparator/ comparator related costs	20	29.0	2	15.4	22	26.8
Drug dosing	15	21.7	2	15.4	17	20.7
Treatment duration	9	13.0	2	15.4	11	13.4
Not considering one/some comparators	9	13.0	2	15.4	11	13.4
Adherence/ compliance	10	14.5	1	7.7	11	13.4
Drug wastage or multiple administration from a via	9	13.0	1	7.7	10	12.2
Drug discontinuation	6	8.7	4	30.8	10	12.2
other BIA critiques	34	49.3	5	38.5	39	47.6
Any BIA critiques	69	100.0	13	100.0	82	100.0