

***Transforming Everyday Life
into Extraordinary Ideas***



Application of Text Data Mining Approaches to the Canadian Longitudinal Study on Aging (CLSA) Medication Dataset: Development and Validation

Benoit Cossette, BPharm, PhD

Lauren Griffith BSc, MSc, PhD

**Lisa Dolovich BScPhm, PharmD,
MSc**

DSECT/DSEN Seminar Series

Nov 22, 2018



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

CLSA Medication Coding Group

Benoit Cossette

Lauren Griffith

Lisa Dolovich

Patrick Emond

Dee Mangin

Jinhui Ma

Medication Coders: Helga Weigelin and Bevonie Brown

**With acknowledgment of Kathryn Nicholson,
Joanne Ho and Kasia Makara**



Overview

- Overview CLSA
- Challenges with medication collection
- Medication cleaning protocol development, results to date, validation, and lessons learned

What is the Canadian Longitudinal Study on Aging (CLSA)?

A research platform – infrastructure to enable state-of-the-art, interdisciplinary population-based *research* and *evidenced-based* decision-making that will lead to better health and quality of life for Canadians.



Canadian Longitudinal Study on Aging (CLSA)

- Strategic initiative of CIHR; on Canadian research agenda since 2001
- More than 160 researchers and collaborators – 26 institutions
- Multidisciplinary – biology, genetics, medicine, psychology, sociology, demography, economics, epidemiology, nutrition, health services
- Largest research platform of its kind in Canada for breadth and depth
- Following 50,000+ Canadians aged 45-85 at baseline for 20 years



CLSA Leads



**Co-principal Investigator
Christina Wolfson (McGill)**



**Lead Principal Investigator
Parminder Raina (McMaster)**



**Co-principal Investigator
Susan Kirkland (Dalhousie)**

CLSA Research Platform

50,000 women and men aged 45 - 85 at baseline

Target: 20,000
Actual: 21,241
Randomly selected within
provinces

Target: 30,000
Actual: 30,097
Randomly selected
within 25-50 km of 11 sites

Questionnaire
• By telephone (CATI)

Questionnaire
• In person, in home (CAPI)

2010 - 2015

2015

2018

Clinical/physical tests
Blood, urine
• @ Data Collection Site

Participants
aged 45 to 85
at baseline
(51,338)

20 Years

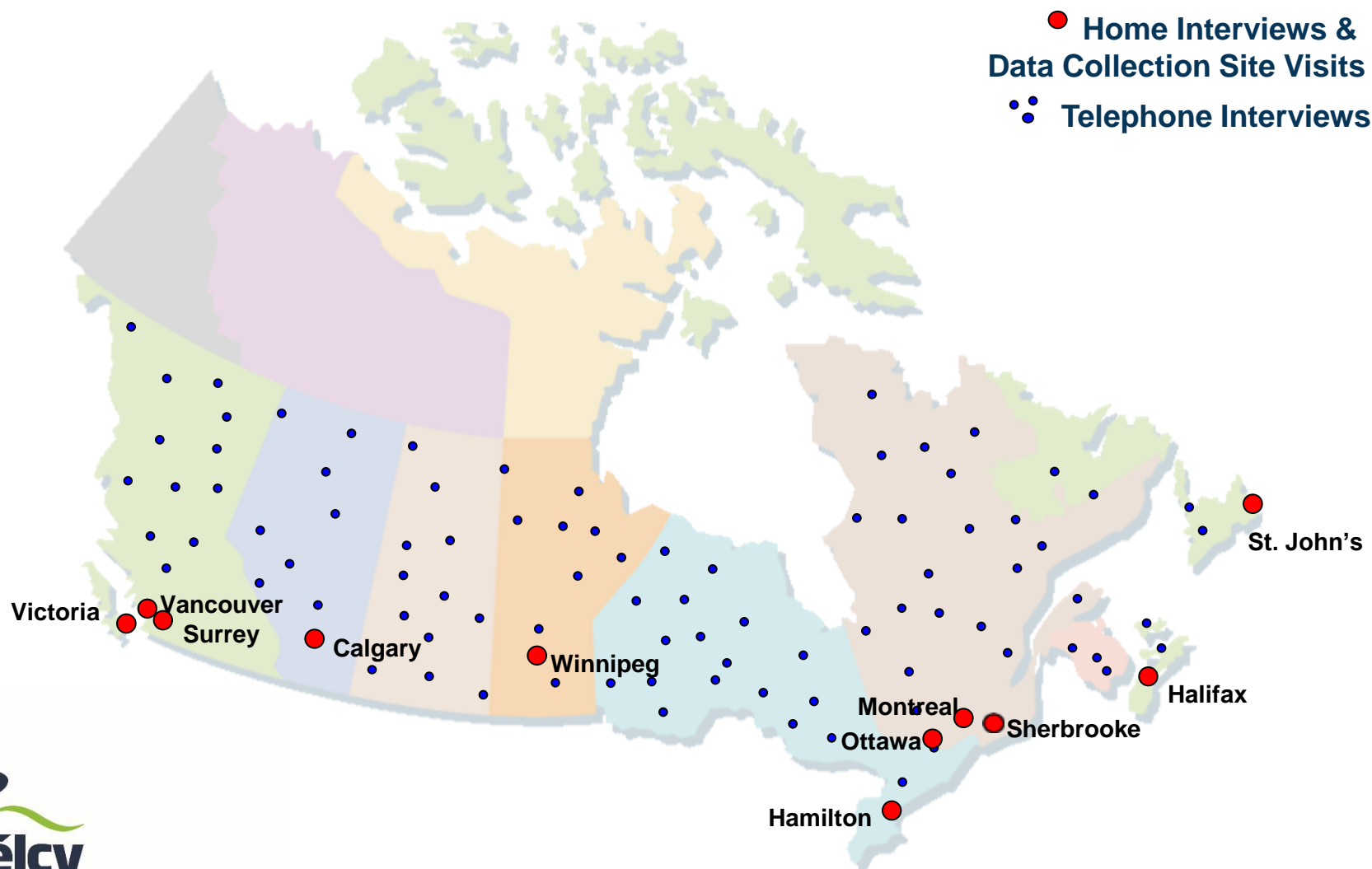
Baseline FU-1 FU-2 FU-3 FU-4 FU-5 FU-6

Active follow-up every 3 years



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

National Scope



Depth and Breadth of CLSA

PHYSICAL & COGNITIVE MEASUREMENTS

- Height & weight
- Waist and hip measurements
- Blood pressure
- Grip strength, timed up-and-go, chair raise, 4-m walk
- Standing balance
- Vision (retinal imaging, tonometer & visual acuity)
- Hearing (audiometer)
- Spirometry
- Body composition (DEXA)
- Bone density (DEXA)
- Aortic calcification (DEXA)
- ECG
- Carotid intima-media thickness (ultrasound)
- Cognitive assessment (30-minute battery)
- Biospecimen collection (blood and urine)

HEALTH INFORMATION

- Chronic disease symptoms (11 chronic conditions)
- Medication and supplement intake & compliance
- Women's health
- Self-reported health-care utilization
- Oral health
- Administrative data linkage health services, drugs and other administrative databases

PSYCHOSOCIAL

- Social participation
- Social networks and support
- Caregiving and care receiving
- Mood, psychological distress
- PTSD
- Injuries and consumer products
- Work-to-retirement transitions
- Personality traits
- Retirement planning
- Social inequalities
- Mobility-lifespace
- Built environments and contextual factors
- Income, wealth and assets

LIFESTYLE & SOCIODEMOGRAPHIC

- Smoking
- Alcohol consumption
- Physical activity (PASE)
- Nutrition (nutrition risk and food frequency)
- Ethnicity/race/gender
- Birth location
- Marital status
- Education

Terminology

- Tracking Cohort
 - Target - 20,000 participants from all 10 provinces, followed through Computer-Assisted Telephone Interviews (60 minutes at baseline)
 - **21,241 recruited***
- Comprehensive Cohort
 - Target - 30,000 participants living within 25 km (or 50 km) of a CLSA Data Collection Site (DCS)
 - Followed through in-home interviews (60 minutes) and physical assessments (2-3 hours) at a DCS
 - **30,097 recruited***

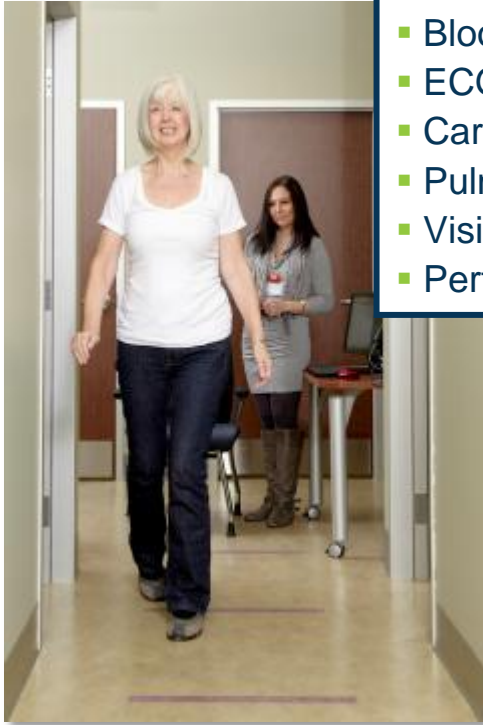


CLSA Data Collection

Data Collection Site

Physical Assessments:

- Height, Weight, BMI
- Bone Density, Body Composition, Aortic Calcification
- Blood Pressure
- ECG
- Carotid Intimal-Medial Thickness
- Pulmonary Function
- Vision & Hearing
- Performance testing



Biospecimen Collection:

- Blood
- Urine



Cognitive Assessments:

- Neuropsychological Battery
 - Memory
 - Executive function
 - Reaction time

Medication Data Collection



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

Entering DIN Numbers & Drug Names

Always use the DIN number field when entering a drug.

Medication 1

Identification of the medication

☒ Drug Identification Number (DIN) 01917056 : ARTHROTEC 50 (MISOPROSTOL 200µG, DICLOFENAC SODIUM 50MG)

☐ Name of Medication

If the DIN number isn't available type the name of the medication in the DIN field, as you are typing, a list of drugs will appear, locate the correct one and click on it to select it.

Medication 3

Identification of the medication

☒ Drug Identification Number (DIN)

☐ Name of Medication

How much, what dose is the medication?

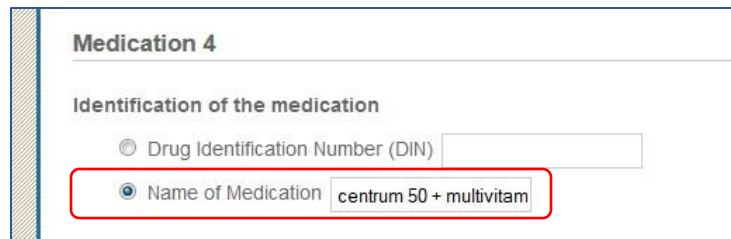
(INTERVIEWER - ENTER ONLY A NUMBER)
(DON'T READ)WHAT ARE THE UNITS?

02260557 : CLARITIN ALLERGY + SINUS EXTRA STRENGTH (LORATADINE 10MG, PSEUDOEPHEDRINE SULFATE 240MG)
02019973 : CLARITIN SYRUP 1MG/ML (LORATADINE 1MG)
00782696 : CLARITIN TAB 10MG (LORATADINE 10MG)
02237734 : CLARITIN (LORATADINE 10MG)
02243688 : CLARITIN EYE ALLERGY RELIEF (OXYMETAZOLINE HYDROCHLORIDE .25MG)
02243690 : CLARITIN ALLERGIC CONGESTION RELIEF (OXYMETAZOLINE HYDROCHLORIDE 0.05%)
01945157 : CLARITIN ALLERGY + SINUS (LORATADINE 5MG, PSEUDOEPHEDRINE SULFATE 120MG)
02241523 : CLARITIN KIDS (LORATADINE 1MG)

Entering DIN Numbers & Drug Names

Multivitamins and herbal supplements will not have DIN's.

Enter the name of the supplement in the “*Name of Medication*” field as shown below



The screenshot shows a form titled "Medication 4". Under the heading "Identification of the medication", there are two radio button options. The first option is "Drug Identification Number (DIN)" with an empty text box next to it. The second option is "Name of Medication" which is selected (indicated by a filled radio button). Next to this option is a text box containing the text "centrum 50 + multivitam". A red rectangular box highlights the "Name of Medication" radio button and its associated text box.

New Question Added (Dec. 2014)

Was this medication prescribed by a doctor or is it a non-prescription medication?

- Yes
- No
- Don't know/No answer
- Refused

NB :This questions was added to try a determine why we have so many “medication” without DINs

Additional Information Collected

- Dosage (quantity, units)
- Frequency (e.g., once a day, once a week)
- Duration (< 6 mo, 6 mo-1 yr, >1 yr)
- Start date
- Reason(s) for use

Health Canada Drug Product Database


The Drug Product Database (DPD) is maintained by Health Canada.

Included drugs:

- Drugs for humans and animals
- Disinfectants
- Radiopharmaceutical drugs
- Biological drugs for humans

Source: <https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/drug-product-database/more-info.html>

Health Canada DPD - online query

**Government of Canada**
Gouvernement du Canada

Franglais

Search Canada.ca

Jobs ▾Immigration ▾Travel ▾Business ▾Benefits ▾Health ▾Taxes ▾More services ▾

Home → Drugs & Health Products → Drug Products → Drug Product Database → Drug Product Database online query

Drug Product Database online query

From [Health Canada](#)

Due to the fact that the information originated with an organization that is not subject to the Official Languages Act, the document may only appear in the language in which it was written. Translations of the document are the responsibility of the sponsor involved.

Search criteria

You may search by **either** a) drug identification number (DIN), b) Anatomical Therapeutic Chemical (ATC) code, or c) by company or one or more of the various other product characteristics listed. When typing inside fields, do not include punctuation marks such as hyphens, commas, colons, brackets and wildcard characters (%).

Search by drug identification number

Drug identification number (DIN):

Search by Anatomical Therapeutic Chemical

Anatomical Therapeutic Chemical (ATC):

Search by other criteria

Search by other criteria

Status:

Company:

Product name:

Active ingredient(s):

Active ingredient group number:

Class(es):
Disinfectant
Human
Radiopharmaceutical
Veterinary

Route(s) of administration:
Intra-Arterial
Intra-Articular
Intrabranhial
Intrabursal
Intracardiac
Intracaudal

Dosage form(s):
0-Unassigned
Aerosol
Aerosol, Foam
Aerosol, Metered Dose
Powder (Oral)

Schedule(s):
CDSA Recommended
Ethical
Homeopathic
Narcotic
Narcotic (CDSA)

SearchReset

Health Canada DPD - online query

Search results summary

From [Health Canada](#)

[New search](#)

Search criteria

- Status: Marketed
- Product name: lipitor
- Class(es): Select all
- Route(s) of administration: Select all
- Dosage form(s): Select all
- Schedule(s): Select all

Search results

Filter items Showing 1 to 4 of 4 entries | Show **10** entries

List of returned drug products

Status	DIN	Company	Product	Class	PM 1	Schedule	# 2	A.I. name 3	Strength
Marketed	02230711	PFIZER CANADA INC	LIPITOR	Human	Yes	Prescription	1	ATORVASTATIN (ATORVASTATIN CALCIUM)	10 MG
Marketed	02230713	PFIZER CANADA INC	LIPITOR	Human	Yes	Prescription	1	ATORVASTATIN (ATORVASTATIN CALCIUM)	20 MG
Marketed	02230714	PFIZER CANADA INC	LIPITOR	Human	Yes	Prescription	1	ATORVASTATIN (ATORVASTATIN CALCIUM)	40 MG
Marketed	02243097	PFIZER CANADA INC	LIPITOR	Human	Yes	Prescription	1	ATORVASTATIN (ATORVASTATIN CALCIUM)	80 MG

Health Canada DPD - extract

1	ingredient	Brand_name	TC_ATC_NUMBER	TC_AHFS_NUMBER	strength	strength_unit	Drug_identification_number
16925	CEFAZOLIN (CEFAZOLIN SODIUM)	CEFAZOLIN FOR INJECTION, USP	J01DB04	08:12.06.04	500	MG	2237137
16926	CEFAZOLIN (CEFAZOLIN SODIUM)	KEFZOL	J01DB04	08:12.06.04	1	G	322296
16927	CEFAZOLIN (CEFAZOLIN SODIUM)	KEFZOL	J01DB04	08:12.06.04	10	G	411450
16928	CEFAZOLIN (CEFAZOLIN SODIUM)	KEFZOL	J01DB04	08:12.06.04	500	MG	322288
16929	CEFAZOLIN (CEFAZOLIN SODIUM)	KEFZOL ADD-VANTAGE INJ 1GM/VIAL	J01DB04	08:12.06.04	1	G	411442
16930	CEFAZOLIN (CEFAZOLIN SODIUM)	KEFZOL ADD-VANTAGE INJ 500MG/VIAL	J01DB04	08:12.06.04	500	MG	411434
16931	CEFAZOLIN SODIUM	CEFAZOLIN FOR INJECTION USP	J01DB04	08:12.06.04	20	G	2318830
16932	CEFEPIME (CEFEPIME HYDROCHLORIDE)	CEFEPIME FOR INJECTION	J01DE01	08:12.06.16	2	G	2319039
16933	CEFEPIME (CEFEPIME HYDROCHLORIDE)	MAXIPIME FOR INJECTION -PWS IV IM 1G	J01DE01	08:12.06.16	1	G	2163632
16934	CEFEPIME (CEFEPIME HYDROCHLORIDE)	MAXIPIME FOR INJECTION-PWS IV 2GM	J01DE01	08:12.06.16	2	G	2163640
16935	CEFIXIME	AURO-CEFIXIME	J01DD08	08:12.06.12	400	MG	2432773
16936	CEFIXIME	SUPRAX	J01DD08	08:12.06.12	100	MG	868965
16937	CEFIXIME	SUPRAX	J01DD08	08:12.06.12	100	MG	2195992
16938	CEFIXIME	SUPRAX	J01DD08	08:12.06.12	200	MG	868973
16939	CEFIXIME	SUPRAX	J01DD08	08:12.06.12	400	MG	868981
16940	CEFIXIME	SUPRAX	J01DD08	08:12.06.12	400	MG	2195984
16941	CEFIXIME	SUPRAX - TAB 200MG	J01DD08	08:12.06.12	200	MG	2195976
16942	CEFOTAXIME (CEFOTAXIME SODIUM)	CEFOTAXIME SODIUM FOR INJECTION, BP	J01DD01	08:12.06.12	1.0	G	2261510
16943	CEFOTAXIME (CEFOTAXIME SODIUM)	CEFOTAXIME SODIUM FOR INJECTION, BP	J01DD01	08:12.06.12	2.0	G	2261529
16944	CEFOTAXIME (CEFOTAXIME SODIUM)	CEFOTAXIME SODIUM FOR INJECTION, BP	J01DD01	08:12.06.12	500	MG	2261502
16945	CEFOTAXIME (CEFOTAXIME SODIUM)	CLAFORAN 0.5GM/VIAL	J01DD01	08:12.06.12	500	MG	546208
16946	CEFOTAXIME (CEFOTAXIME SODIUM)	CLAFORAN 1GM/VIAL	J01DD01	08:12.06.12	1	G	546216
16947	CEFOTAXIME (CEFOTAXIME SODIUM)	CLAFORAN 2GM/VIAL	J01DD01	08:12.06.12	2	G	546224
16948	CEFOTAXIME (CEFOTAXIME SODIUM)	CLAFORAN ADD-VANTAGE VIALS 1GM/VIAL	J01DD01	08:12.06.12	1	G	1989758

Licensed Natural Health Products Database (LNHPD)

Managed by Health Canada and includes information on **licensed** natural health products, including:

- vitamin and mineral supplements
- herb and plant-based remedies
- traditional medicines like Traditional Chinese Medicines or Ayurvedic (Indian) Medicines
- omega 3 and essential fatty acids
- probiotics
- homeopathic medicines
- many everyday consumer products, like certain toothpastes, antiperspirants, shampoos, facial products and mouthwashes

Source: <https://www.canada.ca/en/health-canada/services/drugs-health-products/natural-non-prescription/applications-submissions/product-licensing/licensed-natural-health-products-database.html>



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

Findings from the CLSA baseline evaluation

- 42% of medications were coded by the interviewers using a medication name / DIN proposed by the CLSA software linked to an extract from the Health Canada Drug Product Database (DPD).
- The remaining entries were entered in the '*name of medication*' field in free text.
- Based on a sample of 100 participants, 73% of the free text entries were for natural products.

Start date / duration

Decision to drop the start date

- Start date:
 - Selected from calendar
 - 28% in January
 - 5% to 8% in February – December
- Decision to drop the start date Duration:
 - 6 months to one year
 - More than one year
 - ...

Protocol: Computer algorithm development Objectives

- To describe the process used to develop an original computer generated approach to clean medication name data that have been recorded in the CLSA database.
 - Agreement for same medication name is not the same as agreement for same DIN (DIN match).
- To conduct manual re-coding on the remaining uncleaned medication data to inform the refinement of the computer generated approach.
- To identify common medication data entry errors that can be used to support more accurate recording of medications data in the CLSA data in future stages of data collection.



Protocol: Computer algorithm development Methods

- Computer generated approach was applied to all CLSA input entries that required cleaning. The intent of this application was to match or clean the prescribed medication or natural product number entered during a participant's interview, (an input) to one or more product(s) in the DPD or NHP.
- The approach uses multiple distinct algorithms, known hereafter as a test, for matching inputs to drugs or naturals.
- All tests are ordered sequentially such that once an input is matched it is no longer considered in the remaining tests.

Protocol: Computer algorithm development Methods

- Programming conducted using SQL (a database scripting language) and PHP (a general programming language).
- The Health Canada and the CLSA data repository (Opal) data were loaded into a MySQL database using SQL.
- Some pre-processing was conducted on these databases before using PHP to enhance performance and increased speed of matching.
 - For instance, the "simple" test compares the free text to drug names by ignoring non-alpha-numeric characters.

Results so far: Applying computer based algorithm tests

Algorithms

- **Direct:** matches the brand or product name (not case sensitive)
- **Code:** matches 6 or more consecutive digits which matches a DIN or NPN (with or without leading 0's)
- **Word:** matches the brand or product name as a sub-string of the input
- **Simple:** matches ignoring non alpha-numeric characters
- **No parentheses:** matches ignoring any words found in (parentheses)
- **No-units:** matches ignoring any numerical units

Algorithm - direct

Direct: matches the brand or product name (not case sensitive)

Examples:

Input	Match
acetaminophen 500mg	ACETAMINOPHEN 500MG
acetaminophene	ACETAMINOPHENE
Acid Reducer	ACID REDUCER
actonel	ACTONEL
adalat	ADALAT

Algorithm - code

Code: matches 6 or more consecutive digits which matches a DIN or NPN (with or without leading 0's)

Examples:

Input	Match
00026158 Senokot Sennosides 8.6 mg	Senokot Tablets
02237556 Eur-Fer 300mg	Euro-Fer
02367335 sandoz latanoprost 1 drop	SANDOZ LATANOPROST
02371987 Mar-Atenolol 50mg	MAR-ATENOLOL
02374420 Apo-Anastrozole 1mg	APO-ANASTROZOLE

Algorithm - word

Word: matches the brand or product name as a sub-string of the input

Examples:

Input	Match
adalat XL	ADALAT
acetaminophen extra fort	ACETAMINOPHEN
Alendronate sodium (Fosamax)	ALENDRONATE
Altace HCL	ALTACE
apo metoprolol	METOPROLOL

Algorithms - simple

Simple: matches ignoring non alpha-numeric characters (including spaces)

Examples:

Input	Match
Act-Atorvastatin	ACT ATORVASTATIN
advil liqui gels	ADVIL LIQUI-GELS
apo hydro	APO-HYDRO
Apo- Hydro	APO-HYDRO
apometoprolol sr	APO-METOPROLOL SR

Algorithm – no parentheses

No parentheses: matches ignoring any words found in (parentheses)

Examples:

Input	Match
EVISTA (60MG)	evista
MYLAN-METOPROLOL (TYPE L)	mylan-metoprolol
Vitalux(areds)	Vitalux
Whey Powder (Vanilla flavour/Chocolate Flavour)	Whey powder

Algorithm – no unit

No-units: matches ignoring any numerical units

Examples:

Input	Match
APO-SALVENT 5MG/ML	apo-salvent
ASPIRIN 81MG	aspirin
calcium/magnésium 500 mg	Calcium/Magnesium
synthroid.05 mg	SYNTHROID
synthroid.75mg	SYNTHROID

		drug	%	natural	%	drug+nat	%	total	%	
direct	single	3079	32.8%	1907	20.3%	4986	53.1%	5076	54.1%	
	multiple	90	1.0%	0	0.0%	90	1.0%			
code	single	31	0.3%	18	0.2%	49	0.5%	51	0.5%	
	multiple	2	0.0%	0	0.0%	2	0.0%			
word	single	1066	11.4%	441	4.7%	1507	16.1%	2106	22.4%	
	multiple	282	3.0%	317	3.4%	599	6.4%			
simple	single	97	1.0%	84	0.9%	181	1.9%	224	2.4%	
	multiple	22	0.2%	21	0.2%	43	0.5%			
no-parens	single	4	0.0%	4	0.0%	8	0.1%	12	0.1%	
	multiple	2	0.0%	2	0.0%	4	0.0%			
no-units	single	114	1.2%	22	0.2%	136	1.4%	323	3.4%	
	multiple	186	2.0%	1	0.0%	187	2.0%			
total matched		4669	49.8%	2788	29.7%	7457	79.5%	7457	79.5%	
						total unmatched		1926	20.5%	
						total inputs		9383		
direct:	Matches the brand or product name (not case sensitive)									
code:	Matches 6 or more consecutive digits which matches a DIN or NPN (with or without leading 0's)									
word:	Matches the brand or product name as a sub-string of the input									
simple:	Matches ignoring non alpha-numeric characters									
no-parens:	Matches ignoring any words found in (parentheses)									
no-units:	Matches ignoring any numerical units									
single:	The input matches a single DIN or NPN									
multiple:	The input matches more than one DIN or NPN									

Spelling errors

- Acetaminophen
 - Acetamin**a**phen
 - Acetom**e**n**o**phen
- Actonel
 - Act**a**nel
 - Act**a**nol
 - Acton**a**l
- Timulol (timolol)
- Exium (Nexium)
- Tamsulosin
hydrochlorathiazide
(hydrochloride)
- Apothyazide
 - Apo-triazide
(triamterene /
hydrochlorothiazide)OR
 - Apo-hydro
(Hydrochlorothiazide)

Examples of manual coding decision rules to code for DRUG NAME– based on iterative coding exercises (not all shown)

Coding Issue	Manual Code
EXACT MATCH (ie has DIN there, exact product name, if dosage available this might help)	code exact DIN from DPD/NHPD
BRAND NAME but no dosage	code DIN from DPD that is the lowest number DIN
Generic NAME but multiple manufacturers	code DIN from DPD that is the lowest number DIN
Incorrect Spelling	Put in correct spelling and then choose lowest number DIN
Drug class do not code further	Drug class do not code further
Drug class – code further	Drug class – code further
Not a drug	Not a drug
Not enough information	Not enough information



clsa élcw

Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

Study Two: Validation (in progress)

- To determine the extent of agreement between the original and refined computer generated approach (test) and the manual re-coding (gold standard) for cleaning the medication data in the CLSA database.
- To assess the extent of agreement when comparing the original computer generated approach (test) with the refined computer generated approach (gold standard), as well as when the order of the algorithms are applied in different sequences.
- To assess whether the extent of agreement changes when re-coding prescribed medication vs. natural products or when re-coding among participants with specific chronic diseases.

Sample Size

- Based on 1) the presence of free text medication data; and 2) the presence of at least one prescribed medication or natural product within the alpha-numeric data.

	Agreement	Precision	Confidence Interval	Sample Size
Best Scenario	0.95	0.02	0.95	457
Worst Scenario	0.85	0.03	0.95	545

Lessons learned (so far)

- Manual entry by interviews derives many varied data entry issues (despite training)
- Many entries do not match to DPD or NHPD and so manual coding will always be needed
 - Many international products
 - Complex combination products, esp NHPD, vitamins
 - Formatting such as parentheses, brackets cannot always be dealt with by computer
 - Nondrug products
- If more than 1 manual coder they all need to use the exact same drug/NHP database
- Will always be behind with new products unless DPD/NHPD can be accessed in real time (versus updated in the back end every 3-6 months)

Conclusions

- The process of developing and validating the algorithm is challenging
- A well working computer algorithm will be a great asset to CLSA for retrospective and prospective medication data coding
- A clean CLSA medications database will be a great asset to Canadian researchers

Extra Slides re Data Access



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

Data Access



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

CLSA Data Access

- Designed as a research study, funded as a **research platform**
- Data **available** to researchers and trainees based in academic settings and research institutes in Canada and elsewhere
- **2019 application deadlines:**
 - February 25
 - June 3
 - September 23

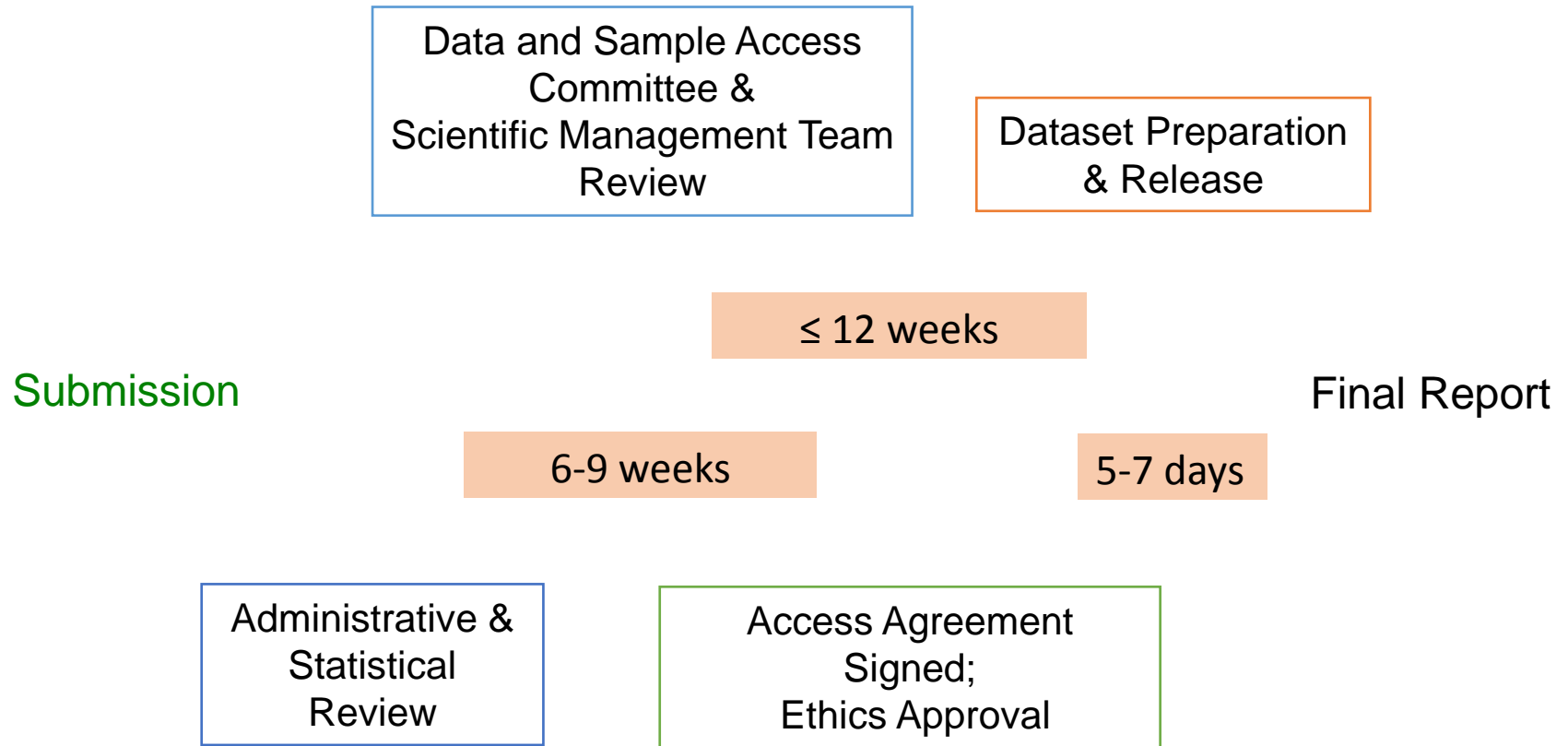


CLSA Data Access Guiding Principles

- The rights, privacy and consent of participants must be protected
- Confidentiality and security of data and biospecimens must be safeguarded
- CLSA data and biospecimens are resources to be used optimally to **support research to benefit all Canadians**



Data Access Timeline



Data Access Fees

- **Partial Cost Recovery Model**

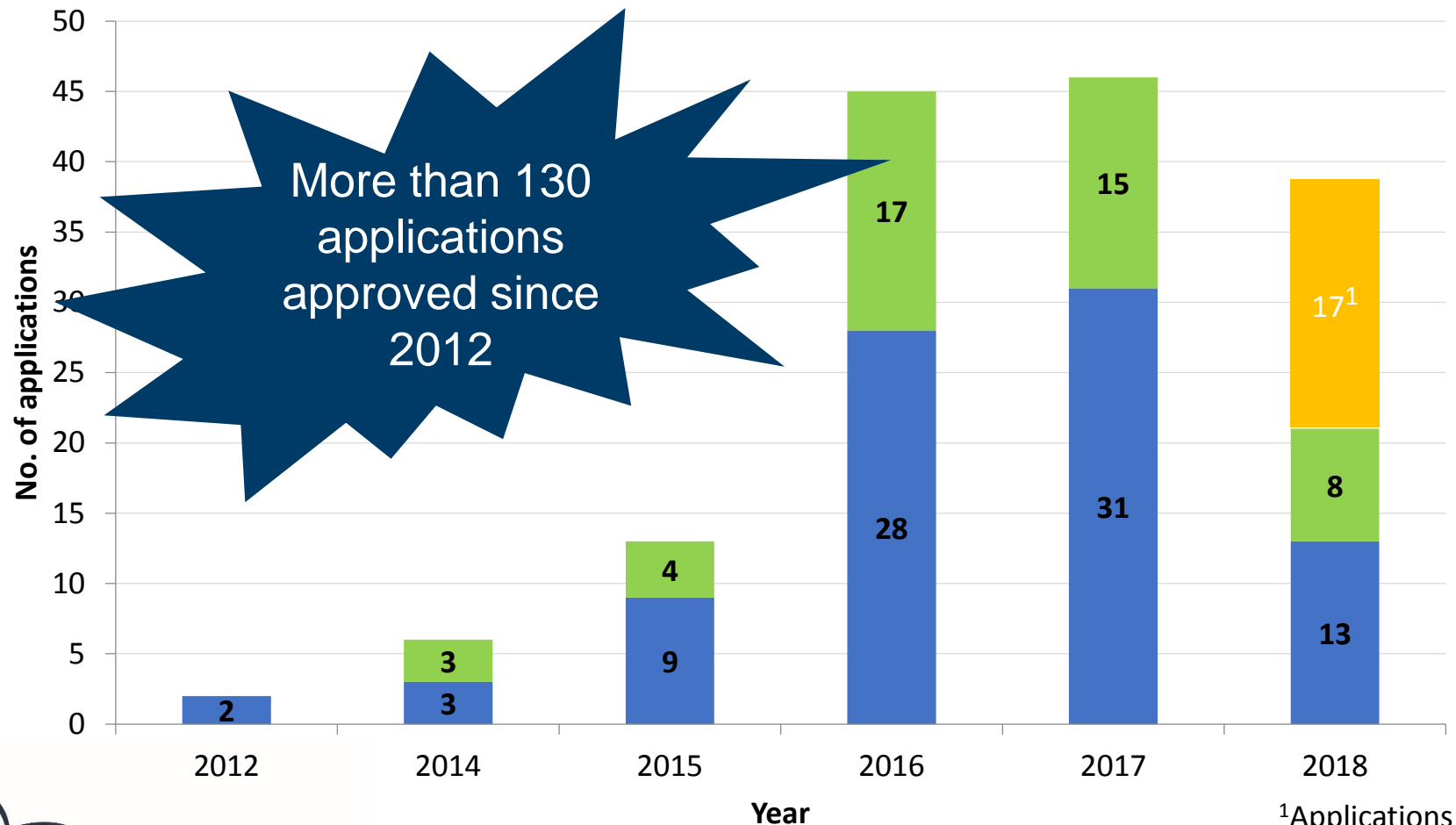
- **Alphanumeric data**

- **\$3,000** for a straightforward alphanumeric dataset
 - **Graduate students** using data solely for thesis research & **Postdoctoral fellows** using data solely for the postdoctoral project are eligible for a **fee-waiver**. Trainees must be enrolled at a Canadian institution or be supported by Canadian funds if working outside Canada.

- **Images & raw data**

- Additional fees of \$1,000 per application are associated with the request for images & raw data.

Approved Applications



Approved Projects Keywords



Resources: www.clsa-elcv.ca



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

Font Size + - Language en fr Search

Update your Contact Details | Contact Us   

About Us | Participants | Partners | Researchers | Data Access | Stay Informed



- [Update your contact details](#)
- [Protecting your privacy](#)
- [Find out how the CLSA platform is being used](#)



- [Spring 2016 data release](#)
- [DataPreview Portal](#)
- [Approved Projects](#)




- [Partners & Supporters](#)
- [Partnering with the CLSA](#)
- [Collaborate and Innovate](#)

clsa élcV

Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement




Information for Researchers & Trainees



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

Font Size + - Language en fr Search

Update your Contact Details | Contact Us



About Us | Participants | Partners | Researchers | Data Access | Stay Informed

Researchers

- Protocols
- Data Collection Tools
- Physical Assessments
- Data Support Documentation
- Approved Project Summaries
- Data Access

Researchers

The CLSA provides documents online to facilitate understanding of the study and how we are gathering and managing the data.

Protocols

- [CLSA Protocol - Executive Summary](#)
- [CLSA Protocol - Full Study Design and Baseline \(2008\)](#)
- [CLSA Protocol - First Follow-up \(2015\)](#)

The protocols listed are based on the applications CLSA submits to CIHR for each funding cycle. As the CLSA data collection progresses, occasionally, some measurements are changed. Updated versions of the protocols will be posted as necessary. Please refer to the Data Collection Tools section to review the specific questions and measurements gathered at each phase of the study.

Data Collection Tools

Over the course of 20 years, the CLSA will be conducting full data collection every three years. At each major data collection event, the questionnaires and physical assessments remain largely the same for consistency, but there will be some additions to the data collection to further enhance the CLSA platform.

[Questionnaires](#)

Physical Assessments

To ensure that physical assessment data are collected, processed, and stored in a consistent, professional, and structured manner at all CLSA sites across the country, Standard Operating Procedures (SOPs) help maintain the integrity of the data collection and data management

Data Preview Portal



DataPreview Portal

SMART TIPS

- Click the 'Help' button on the right to see a step-by-step guide to using the DPP
- Use the main Search Bar on this page to search for predetermined **Areas of Information** or **Scales** only
- For a more detailed search, select 'Variable Properties' under the 'Variable' tab on the left. Expand 'Name' and 'Label' to view search boxes for **Variable Names** and **Variable Labels**

Variable Dataset

- ▶ Variable properties
- ▶ Additional information

▼ Areas of Information

- ▶ Socio-demographic and economic characteristics
- ▶ Lifestyle and health behaviours
- ▶ Health status and functional limitations
- ▶ Diseases
- ▶ Symptoms and signs
- ▶ Medication and supplements

All ▼ e.g. Psychological distress and emotions, Satisfaction with Life Scale 🔍 ? Help

Clear



CMCQ | COM | TMCQ | TRM ▼ ✕

+

Satisfaction with Life Scale ▼ ✕

Variables (32)

Datasets (2)

10 ▼

« < 1 2 3 ... > »

1 - 10 of 32

Name	Label	Dataset
SLS_CONDNEG_COM	SWLS scale: Disagree life conditions excellent	COM
SLS_CONDNEG_TRM	SWLS scale: Disagree life conditions excellent	TRM
SLS_CONDPOS_COM	SWLS scale: Agree life conditions excellent	COM
SLS_CONDPOS_TRM	SWLS scale: Agree life conditions excellent	TRM
SLS_COND_COM	SWLS scale: Life conditions excellent	COM



Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

FAQs

[About Us](#) | [Participants](#) | [Partners](#) | [Researchers](#) | [Data Access](#) | [Stay Informed](#)

Data Access

[DataPreview Portal](#)

[Data Release Timelines](#)

[Data Access Application Process](#)

[Data Access Application Documents](#)

[Data and Biospecimens](#)

FAQs

- › [Data Access Questions](#)
- › [DataPreview Portal Questions](#)
- › [Application Questions](#)

FAQs

Data Access

How do I get access to the data?

Which data formats are available?

What do I do if I would like to obtain biospecimens?

What if there appears to be an error or omission in my data?

Can I apply for data as an international researcher?

What are the fees for access to CLSA data?