

# Identifying a Cohort of People Who Are Transgender and Gender-Diverse Within Saskatchewan's Administrative Health Databases

Gwen Rose<sup>1</sup> , Seanna Goalen<sup>2</sup>, Megan Clark<sup>3</sup>   
and Stéphanie Madill<sup>1</sup> 

<sup>1</sup>College of Rehabilitation Sciences, University of Saskatchewan, Saskatoon, SK, Canada.

<sup>2</sup>Saskatchewan Ministry of Health, Regina, SK, Canada. <sup>3</sup>Academic Family Medicine, University of Saskatchewan, Regina, SK, Canada.

Health Services Insights  
Volume 17: 1–8  
© The Author(s) 2024  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/11786329231222122



**ABSTRACT:** This was a retrospective cohort study. Algorithms were developed to identify a cohort of people who were trans and gender diverse (PTGD) among provincial-level administrative health databases (physician, hospital, emergency department, and pharmacy) from April 1, 2012 to September 30, 2020. Then, healthcare usage was compared between the identified cohort and the general population. There were 6466 unique individuals identified in the cohort, out of a total population of 1.2 million Saskatchewan residents (~0.5%). They had a mean age of 42.5 (SD 17.7) years. 1946 (30.1%) had a female sex marker and 4560 (69.9%) had a male sex marker, which may not indicate their lived gender. The cohort had increased healthcare usage 2 years prior to their index date, compared to the general population, which continued to rise to 1 year past their index date across physician, emergency department visits, and hospitalizations. The results for drugs were mixed. The percentage of PTGD identified in Saskatchewan was comparable to other studies. Healthcare utilization among the cohort was higher than the general population. Further research could use external data sources to validate and improve the cohort identification methods. The large majority of individuals with a male sex marker deserves further investigation.

**RECEIVED:** April 18, 2023. **ACCEPTED:** October 11, 2023.

**TYPE:** Original Research

**FUNDING:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a Sprout grant (#5207) jointly funded by the Saskatchewan Health Research Foundation and the Saskatchewan Centre for Patient-Oriented Research.

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHOR:** Stéphanie Madill, University of Saskatchewan College of Medicine, Suite 3400, 104 Clinic Place, Saskatoon, SK S7N 2Z4, Canada. Email: stephanie.madill@usask.ca

## Background

There have been no attempts to identify people who are transgender and gender-diverse (PTGD) within Saskatchewan's health data on a large scale or to measure their healthcare needs, and there are limited data from many other Canadian regions. High quality data on the health needs of PTGD is required to assess these needs, and to evaluate the effect of interventions and health policies.<sup>1</sup> PTGD have higher needs for mental healthcare, need ongoing cancer screening for the anatomy they have, need ongoing medical supervision for hormone therapy, may need a variety of gender-affirming surgical procedures and are more likely to be subject to violence. This is an issue both because this population is historically underserved by the medical community, but also because most Canadian research regarding the health and healthcare needs of PTGD is from Ontario, a more populous, more urban and centrally-located province.<sup>2</sup>

When this study was designed, we could find no similar examples in the literature. Since completing this project, we have identified 4 studies with similar methodologies that confirm our algorithm.<sup>3–6</sup> Jasuja et al identified 27 227 PTGD out of 74 million adults (0.04%) enrolled in the OptumLabs Data Warehouse from 2006 through 2017, which includes de-identified medical claims data for commercially insured and Medicare Advantage enrollees across the United States. They identified individuals whose data contained at least one International Classification of Diseases, Ninth or Tenth editions (ICD-9/10) diagnostic code for gender identity disorder

(GID), plus one of the following: endocrine disorder not otherwise specified codes, common procedural terminology (CPT) or ICD-9/10 transgender-related procedure codes, or sex-discordant hormone therapy. If an individual's data did not contain a GID code, a combination of any 2 of the other elements was required to validate TGD status. Results indicated that these additional methods identified 31% of the total TGD dataset that would not have been identified using GID codes alone.<sup>3</sup>

Similarly, Quinn et al identified 6456 PTGD out of the 127 608 people (5%) in Kaiser Permanente's insurance database of California and Georgia residents, between 2006 and 2014. Their study used a combination of ICD-9 diagnostic codes, Kaiser Permanente's internal codes and keyword searches within providers' free-text notes to identify their cohort. Keywords identified cohort members, based on natal sex (eg, testes, ovary), hormonal therapy (eg, estrogen, android), and procedures (eg, feminization, vaginectomy). Cases identified through the keyword search were validated by 2 reviewers who examined short text strings that included the keyword; cases without keywords but with at least 2 eligible ICD-9 codes, or the same eligible ICD-9 code used at least twice were also included. The combination of both keywords and ICD-9 codes had the strongest (98%) positive predictive value.<sup>4</sup>

The purpose of this study was to identify a cohort of PTGD from Saskatchewan administrative health databases, accessed through the Saskatchewan Health Quality Council (HQC), then to compare healthcare usage among PTGD with the



overall population. Using electronic databases to identify a cohort of PTGD within a whole province's population has the potential to provide the necessary data for healthcare planning that meets the specific needs of PTGD, such as using emergency room visit data to plan and implement trans-specific mental health services; tracking cancer screening in physician visit data to inform strategies to improve screening rates among PTGD; tracking patterns in hormone prescriptions by the specialty of the provider to assess if training family medicine residents in trans care has shifted hormone prescribing from endocrinologists to family physicians, using physician visit and hospital data in combination to assess wait times for gender affirming surgery and surgical outcomes, and using emergency room data to plan public education interventions to decrease anti-trans violence.

## Methods

### Data

This study involved a retrospective analysis of several population-level Saskatchewan administrative health databases. Linkage and data analysis were performed at the Saskatchewan HQC. These databases were the Person Health Registration System (PHRS), Discharge Abstract Database (DAD), National Ambulatory Care Reporting System (NACRS), physician billing claims data, and Prescription Drug Plan Database (PDPD). The databases are maintained by eHealth and the Saskatchewan Ministry of Health and were accessed through the Saskatchewan HQC (<https://www.saskhealthquality.ca/>). All records for the study period April 1, 2012 to September 30, 2020 were included. Ethics approval for secondary use of health data was received from the University of Saskatchewan Biomedical Research Ethics Board: Bio 2245, and this study meaningfully involved PTGD at all stages, from study design, through data analysis to knowledge dissemination.

### Measures

**Dependent Variable:** Possible PTGD were identified if they met any one of the criteria for any of the health databases on Table 1. For each database, a person's index date was the first occurrence that met a criterion in the study period. Each person's overall index date was the earliest of the database-specific index dates.

A person remained in the cohort if they had continuous Saskatchewan provincial health insurance coverage from 2-years pre-index date until the earliest of the following dates (allowing for gaps of up to 30 days):

- two-years post-index date;
- date of death; or,
- study end date (September 30, 2020).

**Independent Variables:** Independent variables of interest were the types of healthcare service described in the datasets on

Table 1 and time relative to the index date: 2-years pre-index to 2-years post-index.

### Analysis strategy

Health care utilization data for the cohort included all hospitalizations, emergency department visits, physician visits, and drug dispensations in Saskatchewan during the 2-years pre- and 2- years post-index date, as well as any available out-of-province inpatient hospitalizations. These data were then summarized for the whole cohort for the pre- and post-index date periods. For each type of service (eg, hospitalization) and time-period (eg, "1-year pre-index"), the number of individuals in the cohort who had at least once instance of that service in that time-period was determined, and the mean number of services for those individuals was calculated. The means for "1-year pre-index," "1-year post-index" and "2-years post-index" were then compared to the mean for "2-year pre-index" using *t*-tests. Results were considered to be significantly different ( $P < .05$ ) if their 95% confidence intervals did not overlap. The means were also compared to the same healthcare utilization means for all Saskatchewan residents in fiscal year 2019-20. These analyses were completed by SG.

## Results

### Cohort

Applying the cohort criteria resulted in an initial cohort of 7775 unique individuals (See Table 2 for a breakdown of the cohort). Total cohort size decreased to 6466 after excluding cases with a male gender marker and prescriptions for erectile dysfunction drugs. This cohort equates to approximately 0.5% of Saskatchewan's population (1 204 858 in 2021).<sup>7</sup>

Cohort descriptive characteristics are presented on Table 3. The mean age of the cohort was 42.5 years (SD  $\pm$  17.7 years) at index. Thirty percent had sex marked as female and 70% as male. Please note that the sex value is what was recorded in the source database and does not necessarily represent the individual's identified gender.

Seventy-five percent of individuals remained in the cohort for the full 2-years post-index date. The remaining 25% either died (1.4%), or the study end date was reached (23.4%) before the end of 2 years.

### Health care utilization

Table 4 summarizes the cohort's hospitalizations, emergency department visits, physician visits, and drug dispensations. We found that hospitalizations increased over time, both in total number and the number of unique individuals with at least one hospitalization. Emergency room visits and visits to family physicians, referred to as general practitioners in the database, nurse practitioners, and specialists, also all rose significantly from 2-years before to 1-year after the index date. The mean number of beneficiary drugs dispensed rose over

**Table 1.** Cohort inclusion criteria.

DATASET	CRITERIA
Person Health Registration System (PHRS)	Persons where the sex variable changed from “M” to “F” or from “F” to “M,” and age was greater than or equal to one (sometimes incorrect information is recorded when the Health Services Number is assigned before the official Vital Statistics birth registration).
Discharge Abstract Database (DAD) □ hospital data	Inpatient and day surgery stays in Saskatchewan hospitals, and inpatient stays in out-of-province public hospitals, where: <ul style="list-style-type: none"> <li>one of the following ICD-10-CA codes (or their more-detailed sub-codes) was reported in any of the 25 diagnosis fields on the abstract: <ul style="list-style-type: none"> <li>F64 “Gender identity disorders”</li> <li>Z41.80 “Procedures for transgender reassignment”</li> </ul> </li> <li>or</li> <li>one of the following CCI codes (or their more-detailed sub-codes) was reported in any of the 20 intervention fields on the abstract <ul style="list-style-type: none"> <li>1.QE.84 “Construction or reconstruction, penis”</li> <li>1.QE.87 “Excision partial, penis”</li> <li>1.RS.84 “Construction or reconstruction, vagina”</li> </ul> </li> </ul>
National Ambulatory Care Reporting System (NACRS) □ emergency department data	Emergency department visits in Saskatchewan hospitals, where main problem was ICD-10-CA code F64 “Gender identity disorders” (or its more detailed sub-codes). Note: only hospitals in Saskatoon and Regina were reporting to NACRS at the beginning of the study. Prince Albert Victoria Hospital began reporting in 2014. Hospitals in Lloydminster, Maidstone, Meadow Lake, North Battleford and Turtleford began reporting at various times between 2016 and 2018, and the remaining emergency departments in the province began reporting to NACRS after March 31, 2020.
Physician billing claims data	Physician and nurse practitioner billing claims in Saskatchewan where reported ICD-9 diagnosis was 302 “Sexual and gender identity disorders,” and <ul style="list-style-type: none"> <li>fee code was for a general primary care visit to a family doctor or nurse practitioner (5B, 55B, 40B, 41B, 9B)</li> <li>or</li> <li>fee code was for mastectomy from a plastic surgeon or hysterectomy from a gynecologist (120P, 122P, 123P, 124P, 125P, 126P, 84L, 85L, 87L, 88L, 89L).</li> </ul> Because ICD-9 code 302 includes erectile dysfunction, the query was refined to exclude those individuals who had been prescribed any of the following drugs: <ul style="list-style-type: none"> <li>Tadalafil (DINs 02248088, 02248089, 02296888, 02296896)</li> <li>Vardenafil (DINs 02250462, 02250470, 02250489)</li> <li>Sildenafil (DINs 02279401, 02469669)</li> </ul>
Prescription Drug Plan Database (PDPD) □ drugs dispensed	Drug dispensations in Saskatchewan where: <ul style="list-style-type: none"> <li>drug dispensed was Leuprolide (DINs 836273, 2239834)</li> <li>or</li> <li>recorded sex was “M” and drug dispensed was <ul style="list-style-type: none"> <li>Estradiol (oral) (DINs 2449048, 2225190, 2449056, 2148587, 2449064, 2148595)</li> <li>Estradiol (patch) (DINs 2246967, 2244000)</li> <li>Medroxyprogesterone (DINs 2221284, 2244726, 2221292, 2244727, 30937, 2221306, 2277298, 729973)</li> <li>Micronized progesterone (DINs 2439913, 2166704)</li> <li>Note: Estradiol valerate (injectable) is non-formulary and only available by compounding, so was not included</li> </ul> </li> <li>or</li> <li>recorded sex was “F” and drug dispensed was <ul style="list-style-type: none"> <li>Testosterone (injectable) (DINs 29246, 30783)</li> <li>Testosterone gel (DINs 02245345, 02245346, 02249499)</li> </ul> </li> </ul>

ICD-10-CA (International Classification of Diseases, 10th revision, Canadian version) was developed by the Canadian Institute for Health Information (CIHI) as an enhanced version of ICD-10 for disease, injury, poisoning and cause of death classification in Canada.<sup>6</sup>

the same period, as well as the number of individuals with at least one beneficiary drug dispensed; however, the data were less consistent regarding non-beneficiary drug dispensations. Statistically significant changes are most visible from the 2-years pre-index date period through the 1-year post-index period, because the number of people who are TGD in the cohort remained consistent during this time; the data at 2-years post-index are not representative because only approximately 75% of cohort members reached this threshold.

When compared to the general population in Saskatchewan (data from fiscal year 2019–2020), the cohort who were TGD demonstrated higher healthcare use in terms of hospitalizations, emergency department visits, family doctor/general practitioner or nurse practitioner visits, specialist visits, and total physician visits (See Table 4). Moreover, the cohort who were TGD averaged higher numbers of primary care visits than the general population even 2 years prior to the index date, and their healthcare utilization across all databases, except

**Table 2.** Number of people in the cohort, by index year and database where identified.

INDEX YEAR	DATA SOURCE WHERE PERSON IDENTIFIED FOR COHORT					TOTAL
	PHRS	DAD/NACRS	PHYSICIAN BILLING	PDPD (BENEFICIARY)	PDPD (NON-BENEFICIARY)	
2012	0	20	555	122	77	774
2013	217	33	758	82	54	1144
2014	103	46	636	51	35	871
2015	66	42	613	59	37	817
2016	85	44	581	60	48	818
2017	65	42	636	77	40	860
2018	69	37	751	72	54	983
2019	61	60	864	107	52	1144
2020	53	23	231	34	23	364
Total	719	347	5625	664	420	7775

Abbreviations: PHRS, Person Health Registration System; DAD, Discharge Abstract Database; NACRS, National Ambulatory Care Reporting System; PDPD, Prescription Drug Plan Database.

beneficiary: Benefit prescriptions are adjudicated and accepted for payment through the Drug Plan for benefit drugs, fees and supplies for Saskatchewan beneficiaries. Benefit drugs include Formulary drugs and Exception Drug Status (EDS) drugs where the patient has received approval for coverage of the drug(s) under the EDS program.

non-beneficiary: Non-benefit prescriptions include drugs and medical supplies not eligible for coverage under the Drug Plan. They also include prescriptions provided to Saskatchewan residents that are cov.

**Table 3.** Cohort characteristics (n=6466).

Age	
Mean age (in years)	42.5 ± 17.7
Sex marker <sup>a</sup>	
Female	1946 (30.1%)
Male	4560 (69.9%)
Study end reason	
Died	90 (1.4%)
Two-years post-index date	4861 (75.1%)
Study end date	1515 (23.4%)

<sup>a</sup>Sex marker does not necessarily indicate an individual's identified gender. Unless changed by an individual, it indicates the gender assigned at birth based on external sexual characteristics.

non-beneficiary drug dispensations, continued to rise through to 1 year following the index date.

## Discussion

We identified a cohort of 6466 PTGD in Saskatchewan, approximately 75% of whom had healthcare utilization data through 2-years post-index date. This represents a first step in understanding healthcare needs, reducing barriers and decreasing health disparities among PTGD in the province. This cohort constitutes approximately 0.5% of the general SK population, comparable to other North American estimates.<sup>8,9</sup> This

may be an undercount because not all emergency rooms in the province were reporting to NACRS for the whole study period. However, since the emergency rooms in the 2 largest cities were reporting to NACRS throughout and relatively few individuals were identified in this database, any undercount is likely to be small. Our cohort is larger than the 2530, or 0.28% of the province's population, reported as transgender or nonbinary on the 2021 Canadian Census,<sup>7</sup> which may be due to people not disclosing on the Census. There are many reasons why PTGD choose to not disclose on censuses and surveys, including personal safety or wanting to be seen as their lived gender<sup>10</sup>; however, it remains important to know the true number of PTGD to be able to meet their healthcare needs, as individuals who do not wish to be known as trans in general life may still have trans-related healthcare needs. There was a notable discrepancy in the recorded sex of the cohort, with a much higher percentage being recorded as male, even after excluding those with a male sex marker and a prescription for an erectile dysfunction drug. We made this exclusion because the broad ICD-9 diagnostic code 302, the level of data available from physician billing data, also includes erectile dysfunction. The remaining disparity may not represent the actual gender identities of these individuals. In Saskatchewan, there are significant cost and logistical barriers to changing the sex marker on one's birth certificate, and consequently in the administrative health databases.

These results constitute the first large-scale attempt to generate a cohort of PTGD using only provincial-level administrative health databases. A strength of this approach

**Table 4.** Health care utilization.

	PRE-INDEX		POST-INDEX	
	TWO-YEARS	ONE-YEAR	ONE-YEAR	TWO-YEARS
# in cohort	6466	6466	6421	6376
# who died during study period			45(0.7%)	45 (0.7%)
<b>Hospitalizations</b>				
Total # of hospitalizations	1619	1963	2385	1506
# of individuals with at least 1 hospitalization	1064 (16.5%)	1252 (19.4%)	1480 (23.0%)	965 (15.1%)
Mean # of hospitalizations per person [all SK residents in 2019-20: 1.47]	1.52	1.57 <sup>b</sup>	1.61 <sup>a,b</sup>	1.56 <sup>b</sup>
# of urgent admissions	566	686	931	580
# of elective admissions	1053	1275	1452	924
# of day surgeries	876	1071	1149	777
# of inpatient surgeries	735	884	1223	720
Average length of stay per hospitalization (days)	6.12	5.68	7.34	6.94
<b>Emergency department (ED)</b>				
Total # of ED visits	1789	2827	3080	2215
# of individuals with at least 1 ED visit	876 (13.5%)	1220 (18.9%)	1279 (19.9%)	1005 (15.8%)
Mean # of ED visits [all SK residents in 2019-20: 1.84]	2.04	2.32 <sup>a,b</sup>	2.41 <sup>a,b</sup>	2.20 <sup>b</sup>
<b>Physician billing</b>				
<i>Combined family and specialist physician and nurse practitioner (NP) visits</i>				
Total # of physician and NP visits	53238	65831	74084	51888
# of individuals with at least 1 visit	5660 (87.5%)	5983 (92.5%)	6392 (99.5%)	4858 (76.2%)
Mean # of visits [all SK residents in 2019-20: 8.49]	9.41 <sup>b</sup>	11.00 <sup>a,b</sup>	11.59 <sup>a,b</sup>	10.68 <sup>a,b</sup>
<i>Family physician/general practitioner (GP) and nurse practitioner (NP) visits</i>				
# of GP/NP visits	36657	44603	50911	35015
# of individuals with at least 1 GP/NP visit	5528 (85.5%)	5877 (90.9%)	6326 (98.5%)	4722 (74.1%)
GP/NP visits as % of total visits	69%	68%	69%	67%
Mean # of GP/NP visits [all SK residents in 2019-20: 5.94]	6.63 <sup>b</sup>	7.59 <sup>a,b</sup>	8.05 <sup>a,b</sup>	7.42 <sup>a,b</sup>
<i>Specialist (SP) visits</i>				
# of SP visits	16581	21228	23173	16873
# of individuals with at least 1 SP visit	3325 (51.4%)	3917 (60.6%)	4019 (62.6%)	3082 (48.3%)
SP visits as % of total visits	31%	32%	31%	33%
Mean # of SP visits [all SK residents in 2019-20: 4.95]	4.99	5.42 <sup>a,b</sup>	5.77 <sup>a,b</sup>	5.47 <sup>a,b</sup>
<b>Drugs dispensed (beneficiary)</b>				
Total # of drug dispensations	79416	89037	93124	76814
# of individuals with at least 1 drug dispensed <sup>c</sup>	4404 (68.1%)	4570 (70.7%)	4691 (73.1%)	3747 (58.8%)
Mean # of drug dispensations [all SK residents in 2019-20: 18.85]	18.03 <sup>b</sup>	19.48 <sup>a</sup>	19.85 <sup>a,b</sup>	20.50 <sup>a,b</sup>

(Continued)

Table 4. (Continued)

	PRE-INDEX		POST-INDEX	
	TWO-YEARS	ONE-YEAR	ONE-YEAR	TWO-YEARS
Drugs dispensed (non-beneficiary)				
Total # of drug dispensations	32 632	37 094	43 880	34 305
# of individuals with at least 1 drug dispensed	3225 (49.9%)	3450 (53.4%)	4392 (68.4%)	2947 (46.2%)
Mean # of drug dispensations [all SK residents in 2019-20: 10.42]	10.12	10.75	9.99	11.64 <sup>a,b</sup>

Results are reported as numbers, with the percentage of the cohort given in parentheses where relevant.

<sup>a</sup>Value is statistically different from "2-years pre-index" value (based on 95% confidence intervals).

<sup>b</sup>Value is statistically different from the 2019-20 value for all SK residents (based on 95% confidence intervals).

<sup>c</sup>Per instance of an individual having their prescription filled, not per # of drugs.

beneficiary: Benefit prescriptions are adjudicated and accepted for payment through the Drug Plan for benefit drugs, fees and supplies for Saskatchewan beneficiaries. Benefit drugs include Formulary drugs and Exception Drug Status (EDS) drugs where the patient has received approval for coverage of the drug(s) under the EDS program.

non-beneficiary: Non-benefit prescriptions include drugs and medical supplies not eligible for coverage under the Drug Plan. They also include prescriptions provided to Saskatchewan residents that are covered under Indigenous Services Canada's First Nations and Inuit Health Branch, Department of Veteran Affairs or Worker's Compensation.

is that it includes the healthcare usage data for all publicly insured residents of the province. As this was an initial, proof of concept study, and because the PTGD and the healthcare providers on our team stressed how individual each person's transition is, we chose to keep the inclusion criteria broad. Thomson and Katikireddi have discussed several challenges of using administrative health databases, the most important of which concern consent and confidentiality.<sup>1</sup> These are important because individuals do not provide informed consent for their information to be used and they could be at real risk should they be identified as TGD through the research. These risks are mitigated by working through the HQC, which routinely analyses data from highly identifiable, small populations and has robust policies and procedures in place for ensuring confidentiality.<sup>11</sup> Limitations of this approach result from limitations in the data collected. The physician billing database uses the older ICD-9 codes and does not include the qualifying numbers after the decimal point. This resulted in the inclusion of individuals with other "sexual dysfunction" diagnoses. We excluded males with prescriptions for erectile dysfunction drugs, but the continued imbalance between the number of people with male compared to female sex markers suggests that we may not have been able to identify everyone with other "sexual dysfunction" diagnoses. Others may have been falsely included because of lack of specificity in the prescription drug data and ICD-9/10 codes may have caused individuals prescribed leuprolide for other indications (such as children with precocious puberty or cis women with severe endometriosis) to be included. Being restricted to the information in the databases also made it impossible to validate whether all members of the cohort had TGD identities. These 2 limitations may help explain why our cohort has nearly double the number of people who reported trans or non-binary (the only gender-diverse

options available) identities on the 2021 Canadian Census; however, it is difficult to compare our results to the Census because the Census did not distinguish between transfeminine and transmasculine identities.<sup>12</sup> Conversely, we may have accurately identified PTGD who chose not to self-identify on the Census for various reasons.<sup>10</sup> Alternately, our study may have failed to accurately identify all individuals who are TGD because the databases used do not include information about gender-affirming procedures that are performed privately, or out-of-province. Others may have been excluded because of the limitations in the emergency department visit (NACRS) data discussed above, or because we did not include ICD-9- 259.9, "unspecified endocrine disorder,"<sup>75</sup> as an inclusion criterion; similar studies increased their cohort numbers by including this code,<sup>5,6</sup> but, since the physician billing database does not include the qualifying decimal numbers on ICD-9 codes, it was too vague. Some transfeminine people may have been excluded from the cohort because injectable estradiol, which is only available by special compound in Saskatchewan, was not able to be included in our data. Some transfeminine people with a male gender marker who use erectile dysfunction drugs could have been excluded as well. The study may also have excluded other gender diverse people who are not seeking hormone therapy or trans-related surgical procedures, or because the sex marker "X" has not yet been adopted in Saskatchewan's health databases.<sup>13</sup>

Other studies that have used health records to identify cohorts of PTGD have differed from our study in 2 important ways. First, some of these studies have had access to less representative data than ours. Jasuja et al<sup>3</sup> and Quinn et al<sup>4</sup> used commercial health insurance databases whose members may have differed from the general population in important ways. Specifically, these data may underrepresent PTGD due to their

higher likelihood of lacking private health insurance.<sup>3</sup> Rich et al identified a cohort of PTGD within a provincial database of individuals receiving care for HIV in British Columbia.<sup>6</sup> This cohort may have had a larger proportion of PTGD because it was derived from a public healthcare system, but it may be less comparable because it worked from a population of people receiving care for HIV. The second way that our study differs from others is that we did not have access to external data against which to validate the cohort. Quinn et al had free-text notes from healthcare providers and Rich et al had a provider-reported transgender measure with which to validate their respective TGD cohorts.<sup>4,6</sup> Rich et al achieved a 0.2% false-positive rate based solely on ICD-9/10 codes and exogenous sex hormone prescription use.<sup>6</sup> All retrospective cohort studies are limited by the available data. The GID codes in the DAD and NACRS databases and the sexual and gender identity ICD-9/10 codes have limited efficacy in identifying PTGD, because they are overly general and are not used in a standardized way.<sup>14,15</sup> Our inclusion criteria were similar to those of the studies discussed above; therefore, we believe that we were similarly accurate in identifying our cohort, despite not having external data for validation.

Our study went beyond the other published PTGD-cohorting studies by examining healthcare usage within the cohort over time and compared to the general Saskatchewan population. The cohort's increase in healthcare use over the study period may indicate that the index date coincided with individuals' medical and surgical transitions and their increased need for healthcare. Another explanation is that PTGD experience more health problems in general, since hospital admissions classified as urgent and emergency department visits also increased over the study. It is important to remember that since emergency departments outside the 2 largest cities were not reporting to NACRS throughout the study period, emergency department use was likely undercounted. When compared to the general population in Saskatchewan, the TGD cohort demonstrated higher healthcare usage across all utilization categories, except non-beneficiary drug dispensations, confirming the higher self-reported healthcare needs of PTGD.<sup>16</sup> The comparison period for the general population did overlap with the beginning of the COVID-19 pandemic which led to an overall reduction in healthcare usage; however, COVID restrictions were initiated in Saskatchewan on March 23, 2020 and the comparison period ended on March 31, 2020 so we believe this effect to have been minimal.<sup>17</sup>

Future directions for our research include verifying this cohort by testing it against other databases. Computable phenotypes could be calculated to determine the most precise combination of variables. The pattern of increasing healthcare use among our cohort, both over time and compared to the general population, deserves more investigation (Unfortunately, we are not able to calculate an effect size as we no longer have access to the databases). Once the algorithm has been refined and

validated, it could be used to ensure that PTGD are receiving equitable care that meets their needs in a timely fashion, for example by examining mental healthcare, cancer screening rates and surgery wait times. Since the databases include all residents of the province, it should also be possible to assess how many people do not have a primary care provider, as surveys have found that PTGD are more likely than cisgender Canadians to report that they do not have one.<sup>16</sup> The databases could also be used to look for geographic patterns in the data, that is, areas of the province where PTGD are less able to access care, as our group found in focus groups with PTGD.<sup>18</sup> Interventions could be planned to directly address these situations: for example, training more family medicine residents in trans care, developing cancer screening programs specifically for PTGD and using telehealth to reach people in more remote parts of the province, and then they could be followed over time using the administrative health databases. Building a national cohort and comparing healthcare usage nationally and internationally could also be future steps, to assess the effects of TGD-related health policies. While Canada has a tradition of keeping politics at arm's length from healthcare,<sup>19</sup> recent anti-TGD politics from the United States have been felt in Canada with protests against PTGD in Saskatchewan<sup>20</sup> and limits on children who are TGD's rights at school in New Brunswick,<sup>21</sup> the administrative databases could be used to accurately track the age at which individuals begin their medical transitions and the health effects of increased anti-trans protests, violence and policies, such as increased emergency room visits for self-harm and suicidality, increased visits to primary care providers for mental health reasons and increased numbers of prescriptions for antidepressant and anti-anxiety medications; conversely, the databases might show a decrease in healthcare use as a Canadian survey of PTGD found that they avoided seeking necessary healthcare due to fear of transphobia.<sup>16</sup> The results could be used to compare the effects of policies between provinces as healthcare is under provincial jurisdiction in Canada, and to create interventions to both ensure that healthcare services and providers are prepared to address the particular needs of PTGD and to inform PTGD about where and how to access culturally-safe care.

## Conclusions

This study was able to identify a substantial cohort of people who presumably are trans and gender diverse by using provincial-level administrative health databases. Although we were not able to externally validate the cohort, its methodological similarities to other cohorting studies make us confident in the results. It demonstrates that this type of study can identify many more PTGD than are willing to self-identify on surveys or the Census, which is essential for healthcare planning, service delivery and assessment to reduce healthcare disparities among PTGD. This study also found that the cohort used more healthcare services compared to the general population. Further research is needed to determine

the reasons for this increased healthcare use and the optimal ways to meet these needs. In combination with research that has shown that PTGD experience significant discrimination and stigma from healthcare providers and institutions, and that they frequently have unmet healthcare needs,<sup>22,23</sup> this research shows that PTGD need to be considered in healthcare planning at all levels: from training healthcare providers to provide culturally-safe care to PTGD, to reducing systemic barriers to transition-related healthcare and to eliminating obstacles, such as facilitating name and gender marker changes, from healthcare systems. Using whole population administrative health databases has the potential to provide sound evidence about the healthcare needs of PTGD that can be used to make informed decisions regarding healthcare service planning and evaluation while protecting their privacy and safety.

### Author's Note

The views expressed in the submitted article are the authors' own, and not the official position of the University of Saskatchewan, the Saskatchewan Ministry of Health, the Saskatchewan Health Research Foundation, the Saskatchewan Centre for Patient-Oriented Research, or the Saskatchewan Health Quality Council.


### Acknowledgements


The authors would like to thank Beliz AÇan Osman of the Saskatchewan Health Quality Council for her contributions to this project. With permission from her wife, this manuscript is submitted posthumously on behalf of Seanna Goalen. Seanna provided invaluable contributions to this research at all stages from research planning to editing.


### Author Contributions

Seanna Goalen, Megan Clark and Stéphanie Madill designed the study with Beliz AÇan Osman of the Saskatchewan Health Quality Council; they all contributed to the ethics application. The cohort was created, and its healthcare usage determined, by Beliz AÇan Osman. Seanna Goalen performed the t-tests. All the authors contributed to the data interpretation. Gwen Rose wrote the Introduction and Discussion sections and edited the manuscript. Seanna Goalen wrote the Methods and Results sections. Megan Clark and Stéphanie Madill edited the manuscript and Beliz AÇan Osman contributed to the final wording of the text. With permission from her wife, this manuscript is submitted posthumously on behalf of Seanna Goalen.

### ORCID iDs

Gwen Rose  <https://orcid.org/0000-0002-0438-2953>

Stéphanie Madill  <https://orcid.org/0000-0003-1346-5096>

Megan Clark  <https://orcid.org/0000-0001-9061-1144>

### REFERENCES

1. Thomson RM, Katikireddi SV. Improving the health of trans people: the need for good data. *Lancet Public Health*. 2019;4:e369-e370.
2. Ziegler E, Valaitis R, Carter N, Risdon C, Yost J. Primary care for transgender individuals: a review of the literature reflecting a Canadian perspective. *Sage Open*. 2020;10:1-10.
3. Jasuja GK, de Groot A, Quinn EK, et al. Beyond gender identity disorder diagnoses codes: an examination of additional methods to identify transgender individuals in administrative databases. *Med Care*. 2020;58:903-911.
4. Quinn VP, Nash R, Hunkeler E, et al. Cohort profile: Study of transition, outcomes and gender (STRONG) to assess health status of transgender people. *BMJ Open*. 2017;7:1-13.
5. Guo Y, He X, Lyu T, et al. Developing and validating a computable phenotype for the identification of transgender and gender nonconforming individuals and subgroups. *AMIA Annu Symp Proc*. 2020;2020:514-523.
6. Rich AJ, Poteat T, Kochoorn M, et al. Development of a computable phenotype to identify a transgender sample for health research purposes: a feasibility study in a large linked provincial healthcare administrative cohort in British Columbia, Canada. *BMJ Open*. 2021;11:1-9.
7. Table 98-10-0036-01 Broad age groups and gender: Canada, provinces and territories. Statistics Canada Web site. Accessed April 1, 2023. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=9810003601>
8. Meerwijk EL, Sevelius JM. Transgender population size in the United States: a meta-regression of population-based probability samples. *Am J Public Health*. 2017;107:e1-e8.
9. Flores AR, Herman JL, Gates GJ, Brown TNT. How many adults identify as transgender in the United States? *The Williams Institute*. 2016. Accessed April 1, 2023. <https://williamsinstitute.law.ucla.edu/wp-content/uploads/Trans-Adults-US-Aug-2016.pdf>
10. Doyle DM. Transgender identity: development, management and affirmation. *Curr Opin Psychol*. 2022;48:1-6.
11. Privacy Policy. Saskatchewan Health Quality Council Web site. Accessed May 18, 2023. <https://www.saskhealthquality.ca/hqc-privacy-policy/>
12. Codes and classifications. Canadian Institute for Health Information Web site. Accessed April 1, 2023. <https://www.cihi.ca/en/submit-data-and-view-standards/codes-and-classifications#:~:text=ICD%2D10%2DCA%20classi%EF%AC%81es%20diseases,and%20lifestyle%20and%20psychosocial%20circumstances>
13. Change of sex designation. eHealth Saskatchewan Web site. Accessed April 1, 2023. <https://www.ehealthsask.ca/residents/Pages/Sex-Designation.aspx>
14. Deutsch GJ, Keatley J, Mayer G, Hastings J, Hall A. Electronic medical records and the transgender patient: recommendations from the World Professional Association for Transgender Health EMR Working Group. *J Am Med Inform Assoc*. 2013;20:700-703.
15. Shenoy AM, Kashey N, Brown P, Kudler NR. Better representing transgender patients in the electronic health record. *J Health Care Poor Underserved*. 2014;25:1641-1645.
16. Scheim AI, Coleman T, Lachowsky N, Bauer GR. Health care access among transgender and nonbinary people in Canada, 2019: a cross-sectional survey. *CMAJ Open*. 2021;9:E1213-E1222.
17. COVID-19: Saskatchewan declares state of emergency, imposes additional measures to protect Saskatchewan residents. Saskatchewan Web site. Accessed May 18, 2023. <https://www.saskatchewan.ca/government/news-and-media/2020/march/18/covid-19-state-of-emergency>
18. Rose G, McCarron M, Reid M, et al. An investigation of the healthcare needs of individuals who are transgender or gender-diverse in Saskatchewan. Poster presented at: 27th Scientific Symposium of the World Professional Association for Transgender Health; September 2022; Montreal, Canada.
19. *R. v. Morgentaler*. CanLII 90, 30 (Supreme Court of Canada 1988).
20. Activists respond to transphobic backlash ahead of Saskatchewan Rush Pride game. CBC/Radio-Canada website. Accessed August 11, 2023. <https://www.cbc.ca/news/canada/saskatchewan/activists-respond-transphobic-backlash-saskatchewan-rush-pride-game-1.6755636>
21. As New Brunswick changes its LGBTQ policy in schools, advocates worry it's just the beginning. CBC/Radio-Canada website. Accessed August 11, 2023. <https://www.cbc.ca/news/canada/new-brunswick-trans-lgbtq-higgs-1.6889957#:~:text=New%20Brunswick%27s%20recent%20changes%20to,change%20their%20name%20at%20school>
22. MacKinnon KR, Grace D, Ng SL, Sicchia SR, Ross LE. 'I don't think they thought I was ready': how pre-transition assessments create care inequities for trans people with complex mental health in Canada. *Int J Ment Health*. 2020;49:56-80.
23. Teti M, Kerr S, Bauerband LA, Koegler E, Graves R. A qualitative scoping review of transgender and gender non-conforming people's physical healthcare experiences and needs. *Front Public Health*. 2021;9:598455.