

Federal Court



Cour fédérale

Date: 2023011325

**Docket: T-645-20
T-641-20
T-637-20**

Citation: 2023 FC 55

Ottawa, Ontario, January 13 25, 2023

PRESENT: Mr. Justice Pentney

Dockets: T-641-20, T-645-20

BETWEEN:

PATRICK CAIN

Applicant

and

MINISTER OF HEALTH

Respondent

and

**THE PRIVACY COMMISSIONER OF
CANADA**

Intervener

Docket: T-637-20

BETWEEN:

MOLLY HAYES

Applicant

and

MINISTER OF HEALTH

Respondent

and

**THE PRIVACY COMMISSIONER OF
CANADA**

Intervener

PUBLIC JUDGMENT AND REASONS

(Amended Confidential Version issued on January 25, 2023)

I. **Introduction**

[1] This case concerns a rather unusual question: was Health Canada justified in refusing to release some of the second and third characters of postal codes of individuals licensed to grow medical marijuana under the licensing regime that was in place before it was legalized, as well as the names of some of the cities where such licensed production occurred?

[2] What this case is really about, however, is the balance between the fundamental right to personal privacy and an individual's right to access information held by the government. More particularly, the case raises the question of the appropriate analytical approach to measuring privacy risks in relation to the release of information from structured datasets that contain personal information.

[3] The Respondent, Health Canada, released the first character of the relevant postal codes relating to licenses to grow medical marijuana (either for personal use or as a "designated producer" for someone else), but refused to release more information. It takes the position that there is a serious possibility that this data, when combined with other information that is already available, could lead to the identification of specific individuals.

[4] All parties to this proceeding agree that information that could identify a specific individual who has a medical marijuana license is personal information that is protected from disclosure. The explanation for this is simple: individuals obtained licenses by providing medical information about their health condition to justify their use of medical marijuana, and information about one's health is among the most deeply personal information imaginable.

[5] The parties disagree, however, about whether the information in dispute, namely the second and third characters of some postal codes and the names of some cities, is personal information. Health Canada refused to release this information because of the risk that it could be combined with other information already in the public domain to identify specific individuals. The Information Commissioner, on behalf of the Applicants, disputes this.

[6] The parties also disagree about the degree of effort Health Canada is required to undertake in order to release information in a way that protects personal information.

Health Canada says it should not be required to assess each of the hundreds of postal codes within the relevant datasets in order to determine whether any of them pose little or no risk. The Information Commissioner disagrees, arguing that Health Canada has already created a computer code that can automate this process.

[7] The Privacy Commissioner intervened in this case, but limited his submissions to the proper application of the legal tests to the type of structured datasets involved in this case, and the related question of the appropriate analytical framework to be applied in assessing the privacy risks associated with the release of data from such datasets.

[8] For the reasons that follow, the application is dismissed. I find that Health Canada was justified in refusing to release more information, because of the serious possibility that it could lead to a breach of privacy through the identification of an individual in the datasets. I also find that Health Canada was not required to undertake a more detailed analysis of the risks associated with releasing more information pursuant to its obligation to sever and release as much information as is reasonable.

II. Background

[9] These consolidated applications were brought by the Information Commissioner on behalf of the named Applicants, ~~David~~ Patrick Cain and Molly Hayes, pursuant to subsection 41(1) of the *Access to Information Act*, RSC 1985, c A-1 [ATIA], challenging Health Canada's

refusal to disclose parts of the postal codes and, for Ms. Hayes' request, the names of cities associated with licenses to grow medical marijuana. Before examining the procedural history of the access to information requests filed by Ms. Hayes and Mr. Cain, I will begin with a brief explanation about Canadian postal codes, because they are integral to an understanding of the dispute before the Court.

A. *Postal Codes in Canada*

[10] Canadian postal codes contain six characters, divided into two groups of three. The first three characters are called a Forward Sortation Area (FSA), which identify major geographic divisions in an urban or rural location. The last three characters are called a Local Delivery Unit, which identify the smallest delivery zone within an FSA.

[11] The first character of a postal code represents a postal district. Quebec and Ontario, for example, are divided into three and five postal districts respectively. These provinces have one urban area with a population large enough to have a dedicated postal district represented by a letter ("H" for the Montreal region and "M" for Toronto). By way of contrast, although Nunavut and the Northwest Territories comprise a vast swath of Canada's geography, their populations are so small that they share a single FSA.

[12] The second character of an FSA identifies the area as either urban or rural, with a zero indicating a wide-area rural region and all other digits indicating urban areas. The third character of the FSA represents a specific rural region, an entire medium-sized city, or a section of a major city. For example, the first three characters of the Federal Court's Ottawa postal code (K1A)

indicate that its mailing address is located in downtown Ottawa; the final three characters specify the location with a greater degree of precision.

[13] In Canada, an FSA may refer to a densely populated urban location, or a sparsely populated rural area spread over a large territory. Statistics Canada has published a document entitled “2016 Population and Dwelling Count by FSA” which shows that in 2016 the majority of FSAs had populations above 10,000. This document also reveals the wide disparity in population counts, ranging from zero inhabitants to 139,128.

B. *Molly Hayes’ Request (Court File Number: T-637-20)*

[14] In August 2017, Ms. Hayes made an access to information request to Health Canada for the following information:

List of addresses of all licensed personal production ACMPR grow operations in Canada that have been authorized by Health Canada to possess 244 or more indoor marijuana plants, and/or 95 or more outdoor plants, and/or 35,625 grams or more in storage at any time (emphasis in original).

[15] The requested information was obtained by Health Canada under the regulatory regime for the possession and cultivation of medical marijuana that was in place at the time; namely, the *Access to Cannabis for Medical Purposes Regulations*, SOR/2016-230 [ACMPR or the *Regulations*]. Under this scheme, individuals could apply for licenses to grow their own medical marijuana, either at their place of residence or elsewhere, or they could designate someone else to grow it for them. In order to obtain such a license, the individuals had to provide personal

information about the location where cultivation would occur, as well as medical information justifying their use of medical marijuana.

[16] There had been previous regulations that governed such matters, but this scheme has now been replaced by the *Cannabis Act*, SC 2018, c 16 and the *Cannabis Regulations*, SOR/2018-144.

[17] Health Canada responded to Ms. Hayes' request on October 16, 2017. The relevant record, created by the Controlled Substances and Cannabis Branch from a database it maintained, is a list of 575 addresses, including civic numbers, street names, city and province names, and postal codes. Health Canada's Access to Information and Privacy (ATIP) Division found most of the information to be personal information and thus exempt from disclosure under section 19 of the *ATIA*. The ATIP Division applied severance to the record and only disclosed the province names.

[18] On October 31, 2017, Ms. Hayes registered a complaint with the Office of the Information Commissioner about the refusal to disclose the other information. The Information Commissioner agreed with Health Canada that the subsection 19(1) exemption for personal information applied to the civic numbers, street names, and the last three digits of postal codes, and therefore this information was not to be disclosed. However, the Commissioner asked Health Canada to determine whether additional portions of the postal codes and the city names could be disclosed. Health Canada subsequently agreed to release the first character of the postal code but refused to release any other information, claiming it was "personal information"

because it could lead to identification of the licensed individual when combined with other, previously released information. Health Canada also asserted that it was unreasonable to require it to analyze each FSA separately to determine the risk of re-identification. This is discussed in further detail below.

[19] The Information Commissioner agreed with Health Canada that disclosing city names or complete FSAs for locations with a small population created a risk of identification, but she was not convinced that such a risk arose from the disclosure of city names or FSAs for more populous areas. She also disagreed with Health Canada's assertion that it was not reasonable to ask it to analyze each FSA to determine which could be disclosed.

C. *Mr. Cain's Request (Court File Numbers: T-641-20 and T-645-20)*

[20] In October 2017, Mr. Cain requested access to:

A document in a sortable format, such as .txt, .csv, or .xls, showing the first three characters of the postal codes of personal or designated producers of medical cannabis, or alternatively totals by the first three characters of the postal code, with personal and designated growers broken out from each other.

A document in a sortable format, such as .txt, .csv, or .xls, showing the first three characters of the postal codes of registered users of medical marijuana, or alternatively totals by the first three characters of the postal code.

[21] In response to the first request, the Controlled Substances and Cannabis Branch created two spreadsheets of information related to personal and designated cannabis producers for

medical purposes, listing FSAs and corresponding numbers of registered personal producers (11,100) and registered designated producers (673), respectively.

[22] On the second request, the Branch noted that the term “registered user” was not defined under the scheme. Instead, it created a spreadsheet including the province and FSA for 11,843 individuals, who were licensed to cultivate medical marijuana or had designated someone else to do so on their behalf.

[23] Health Canada’s ATIP unit examined these records, and disclosed the first character of approximately 11,773 FSAs on the first request, and the first character of approximately 11,842 FSAs for the second. As it had done with the Hayes complaint, Health Canada refused to disclose the second and third characters of the FSAs pursuant to subsection 19(1) of the *ATIA*. Mr. Cain complained about the incomplete disclosure to the Information Commissioner.

[24] On May 7, 2020, following an investigation of Mr. Cain’s complaint, the Information Commissioner accepted that disclosure of FSAs with small populations would create a serious possibility of identification of individual producers and users, but was not convinced this was the case for most FSAs because their populations were larger. The Information Commissioner found that Health Canada’s blanket refusal to release more information was not justified, because the risk of re-identification of the designated persons did not meet the legal test, and Health Canada’s refusal to undertake the necessary analysis was not justified.

D. *Health Canada’s Final Decisions*

[25] On January 20, 2020, Health Canada responded to the Information Commissioner's reports on the complaints, indicating it did not intend to implement the Information Commissioner's recommendations to disclose the FSAs and cities.

[26] Health Canada maintained its position that the FSAs and cities were personal information that it was obliged to exempt from disclosure under subsection 19(1) of the *ATIA*, and explained that it would not release the information under any of the discretionary exceptions listed in subsection 19(2). It stated that the release of the second and third characters of the FSAs and/or the names of the cities, when combined with other available information (including details disclosed pursuant to previous access requests) would create a serious possibility that individuals could be identified. Health Canada asserted that because of this risk, the information fell within the definition of "personal information" and was therefore exempt from disclosure.

[27] The Information Commissioner, acting pursuant to paragraph 42(a) of the *ATIA*, launched applications for judicial review of the Health Canada final decisions. By order of the Court dated August 27, 2020, the matters were consolidated.

[28] The Privacy Commissioner was granted leave to intervene in the proceeding, and was granted access to the confidential information that had been filed. The Privacy Commissioner was permitted to file a memorandum of fact and law and to make oral submissions; the other parties were granted a right of reply to both.

E. *Issues and Standard of Review*

[29] Two main issues arise in this case:

1. Is the Minister authorized to refuse disclosure of the records at issue pursuant to subsection 19(1) of the *ATIA*, because they constitute personal information?
2. Did the Minister correctly refuse to further sever the records pursuant to section 25 of the *ATIA*?

[30] The parties largely agree on the issues and the law applicable to these cases, other than one question relating to the standard of review. The primary debate between the parties, and the intervener's main focus concerns the application of the legal principles to the particular situation before the Court, including the proper analytical approach to assessing the risk of releasing information from structured datasets such as the medical marijuana licensing records held by Health Canada.

[31] On the standard of review for the first issue, the law is clear. Pursuant to section 44.1 of the *ATIA*, reviews under section 41 are heard *de novo*, as a new proceeding. This has been described as “stepping into the shoes of the Minister” to determine whether the refusal to disclose is authorized under the law (*Suncor Energy Inc. v Canada-Newfoundland and Labrador Offshore Petroleum Board*, 2021 FC 138 [*Suncor Energy*] at para 68). In reality, this means that the Court “is to reach its own conclusion as to whether the information at issue is exempt from disclosure under subsection 19(1), i.e., it must determine whether the mandatory exemption has been applied correctly” (*Canada (Information Commissioner) v Canada (Public Safety and*

Emergency Preparedness), 2019 FC 1279 [*Public Safety*] at para 40). The burden is on Health Canada to establish that it was authorized to refuse to disclose the information. The parties agree on these points.

[32] As regards the second issue, however, the parties diverge. Prior case law has found that a Minister's exercise of discretion pursuant to subsection 19(2) is subject to review on the reasonableness standard (see, for example, *3430901 Canada Inc. v Canada (Minister of Industry)*, 2001 FCA 254, [2002] 1 FC 421 [*Telezone*]; *Public Safety* at para 41). The Respondent submits that the same approach should be followed to reviewing its decision not to sever any of the records pursuant to section 25 of the *ATIA*, citing *Attaran v Canada (National Defence)*, 2011 FC 664 [*Attaran*] at paras 18-19.

[33] The Information Commissioner disagrees with this position, asserting that under section 25 of the *ATIA*, severance is mandatory and it should therefore be treated as part of the *de novo* review power under section 44.1. On this view, the Minister is not authorized to refuse to disclose parts of the record that can reasonably be severed, and therefore the Court should reach its own conclusion on whether section 25 was properly applied to the records. The Information Commissioner submits that *Attaran* should not be followed because it has been overtaken by more recent case law.

[34] It does not appear that this specific question has been addressed in any cases since *Attaran*. The main question on this issue is the degree of effort required to meet the obligation under section 25; flowing from that is the question of whether Health Canada's refusal to sever

and release more information from the records is in accordance with the standard that the law requires.

[35] Two things are not in dispute: (i) parts of the records that are responsive to the requests contain personal information; and (ii) Health Canada has exercised its authority under section 25 to partially sever the records, by releasing the first character of the relevant postal codes. The argument in this case focuses on whether Health Canada is required to undertake a more rigorous review, and whether more information should be disclosed.

[36] Based on this, it is clear that at least part of the records in question contain personal information (e.g. the specific home addresses of licensed users, or the full FSAs for areas with a very small population), and thus Health Canada was authorized to refuse to disclose this portion of the record. No one questions that. The only argument relates to whether more information should have been disclosed. On this point, the debate is about whether conducting the sort of “mosaic analysis” (or linking analysis) that would be required to assess the specific risks associated with releasing the names of some cities, and the second and/or third letters of each FSA, goes beyond what the statute requires of a government institution.

[37] In some respects, the facts in *Attaran* are similar to this case insofar as the records contained some personal information, the government institution had released portions of the requested documents, and the debate concerned whether the government institution’s refusal to release more of the details was justified. Justice Barnes’ analysis on this point is worth citing in full:

[18] I accept that the issue of “whether severability has been duly considered” is to be assessed on the standard of correctness: see *3430901 Canada Inc. v. Canada (Minister of Industry)*, 2001 FCA 254, [2002] 1 FC 421 at para 39 [*Telezone*]. I do not agree, though, that the application of that obligation to the evidence is to be judged on that same basis. In my view, deciding whether photographs are severable is an exercise which requires the application of some professional judgment, and thus the standard of reasonableness applies. Notwithstanding the Court’s obligation to pay deference to the decision-maker’s approach to redaction, I am satisfied that the reasonableness standard is sufficiently robust to deal with situations of clearly unwarranted overreaching by the government.

[19] Section 49 of the ATIA deals with the judicial review of withholding decisions made, *inter alia*, under s 19 of that Act. In *Canada (Information Commissioner) v Canada (Commissioner of the Royal Canadian Mounted Police)*, 2003 SCC 8, [2003] 1 SCR 66, the Supreme Court of Canada carried out a detailed standard of review analysis in connection with this provision and held that the determination of what was or was not “personal information” under s 19 of the ATIA should be reviewed for correctness and that the burden of the proof on that point rests with the government. Once it is determined that the decision-maker has correctly exercised that authority, the Court held that the *de novo* review power is “exhausted”. I take that to mean that in the subsequent assessment of a possible redaction of a record authorized by s 25 of the ATIA or in the balancing of privacy rights against the public interest authorized by ss 8(2)(m)(i) of the *Privacy Act*, the decision-maker’s discretion is reviewable on the standard of reasonableness: see *Attaran v Canada*, 2009 FC 339, 342 FTR 82 at paras 28-32 and *Telezone*, above, at para 47. It follows that the Respondent’s decisions not to redact the detainee photographs and to refuse the release the photographs on public interest grounds are reviewable on the basis of reasonableness.

[38] On this approach, the inexorable consequence of the fact that the Respondent was authorized to refuse to disclose at least part of the records in question, and that it specifically considered which parts of the record could be severed, is that the *de novo* review power is exhausted and the question of whether more information should be disclosed must be reviewed on the standard of reasonableness.

[39] The key question is whether the ruling in *Attaran* has been affected by the Supreme Court of Canada's decision in *Merck Frosst Canada Ltd. v Canada (Health)*, 2012 SCC 3

[*Merck Frosst*].

[40] In *Merck Frosst*, the Supreme Court of Canada set out the key principles that govern the application of section 25:

[237] The heart of the s. 25 exercise is determining when material subject to the disclosure obligation “can reasonably be severed” from exempt material. In my view, this involves both a semantic and a cost-benefit analysis. The semantic analysis is concerned with whether what is left after excising exempted material has any meaning. If it does not, then the severance is not reasonable. As the Federal Court of Appeal put it in *Blank v. Canada (Minister of the Environment)*, 2007 FCA 289, 368 N.R. 279, at para. 7, “those parts which are not exempt continue to be subject to disclosure if disclosure is meaningful”. The cost-benefit analysis considers whether the effort of redaction by the government institution is justified by the benefits of severing and disclosing the remaining information. Even where the severed text is not completely devoid of meaning, severance will be reasonable only if disclosure of the unexcised portions of the record would reasonably fulfill the purposes of the Act. Where severance leaves only “[d]isconnected snippets of releasable information”, disclosure of that type of information does not fulfill the purpose of the Act and severance is not reasonable: *Canada (Information Commissioner) v. Canada (Solicitor General)*, 1988 CanLII 9396 (FC), [1988] 3 F.C. 551 (T.D.), at pp. 558-59; *SNC-Lavalin Inc.*, at para. 48. As Jerome A.C.J. put it in *Montana Band of Indians v. Canada (Minister of Indian and Northern Affairs)*, 1988 CanLII 9466 (FC), [1989] 1 F.C. 143 (T.D.):

To attempt to comply with section 25 would result in the release of an entirely blacked-out document with, at most, two or three lines showing. Without the context of the rest of the statement, such information would be worthless. The effort such severance would require on the part of the Department is not reasonably proportionate to the

quality of access it would provide. [Emphasis added; pp. 160-61.]

[238] That said, one must not lose sight of the purpose of s. 25. It aims to facilitate access to the most information reasonably possible while giving effect to the limited and specific exemptions set out in the Act: *Ontario (Public Safety and Security)*, at para. 67.

[41] The Court found that the role of a reviewing judge is to “consider whether the institutional head had properly applied s. 25.” (*Merck Frosst* at para 232).

[42] In light of *Merck Frosst*, I am persuaded that the question of how much effort is required to meet the section 25 severance obligation should be treated as part of the *de novo* review, rather than as a discretionary decision. The text of the provision and its place in the scheme of the *ATIA* support that view. In my view, the discussion of the standard of review in *Attaran* has been overtaken by the Supreme Court of Canada’s decision in *Merck Frosst*, which is binding on me. Following its guidance, I will examine the Respondent’s decision on severance *de novo*, to determine whether it properly applied section 25.

[43] This approach is also consistent with the nature of the review as required under the Act. Subsection 48(1) sets out the burden of proof in proceedings under section 41, which involves establishing either that the head of the government institution “is authorized to refuse to disclose a record...or a part of such a record...” [emphasis added], and this is mirrored by section 49, which sets out the powers of the Court. Under the interpretation of section 25 adopted in *Dagg v Canada (Minister of Finance)*, 1997 CanLII 358 (SCC), [1997] 2 SCR 403 [*Dagg*] and *Merck Frosst*, the government institution is not “authorized” to refuse to disclose a part of a record that can reasonably be severed under section 25.

[44] To be clear, I understand this to entail a two-step process. First, a reviewing court must determine whether the government institution fulfilled its obligation to consider severance under section 25. If not, the reviewing court must do that, in the context of its *de novo* review. This has been done by the court in many previous cases, and in some of these, the court ordered that portions of records be disclosed as required under section 25 (see, for example, *Concord Premium Meats Ltd. v Canada (Food Inspection Agency)*, 2020 FC 1166 [*Concord*]).

[45] The second step involves assessing *de novo* whether it is reasonable to disclose only a portion of the record. The analysis of this question turns on a number of considerations set out in the case law. The guiding principle has been described in a pithy way by Associate Chief Justice Jerome: “Disconnected snippets of releasable information taken from an otherwise exempt passage are not reasonably severable” (*Canada (Information Commissioner) v Canada (Solicitor General)*), [1988] 3 FC 551 (TD) at page 559, cited with approval in *Merck Frosst* at para 237).

[46] As discussed below, the crux of the issue in this case concerns the second step because the Respondent claims that requiring it to undertake a more detailed analysis of the risks of re-identification goes beyond what is reasonable as contemplated by section 25.

III. Analysis

[47] The parties acknowledge that this is a matter of first impression because none of the prior cases have dealt with the precise questions raised here. Before discussing the specific issues raised by this case, it will be helpful to set out the legal framework that applies.

A. *The Legal Framework*

[48] The key elements of the legal framework that governs this proceeding were summarized by Justice McHaffie in *Public Safety* at paras 25-37, and it is not necessary to repeat them in detail. The points most relevant to this case include the following:

- Both access to information and the protection of privacy have been recognized as fundamental rights. The *ATIA* and the *Privacy Act* have been described as quasi-constitutional by virtue of the rights they seek to protect: *Canada (Information Commissioner) v Canada (Minister of National Defence)*, 2011 SCC 25 at para 40; *H.J. Heinz of Canada Ltd. v Canada (Attorney General)*, 2006 SCC 13 [*Heinz*] at para 28;
- The *ATIA* sets out the general principle that the public has the right to access information in records that are under the control of government, which enhances accountability and transparency in government and promotes an open and democratic society: *Merck Frosst* at paras 1, 21-22;
- The protection of privacy is also a fundamental value, enshrined in the *Canadian Charter of Rights and Freedoms* (for example, the guarantee against unreasonable search and seizure in section 8) as well as the *Privacy Act*. The *Privacy Act* protects “personal information” from release, which is defined in a non-exhaustive and non-restrictive manner. The general definition states: “personal information means

information about an identifiable individual that is recorded in any form” (emphasis added);

- The *ATIA* and the *Privacy Act* must be interpreted in parallel, and since both statutes contain an express exemption of personal information from disclosure, privacy rights must be recognized as “paramount” over access to information to the extent that the information falls within the definition of “personal information”: see *Dagg* at para 48; *Heinz* at para 28;
- The definition of “personal information” in section 3 of the *Privacy Act* is “undeniably expansive” and “deliberately broad”; its intention is to capture “any information about a specific person subject only to specific exceptions” (*Dagg* at paras 65, 68-69);
- Subsection 19(1) of the *ATIA* sets out a mandatory exemption from the right to access for personal information, as defined in section 3 of the *Privacy Act*:

Personal information

19 (1) Subject to subsection (2), the head of a government institution shall refuse to disclose any record requested under this Part that contains personal information.

Renseignements personnels

19 (1) Sous réserve du paragraphe (2), le responsable d’une institution fédérale est tenu de refuser la communication de documents contenant des renseignements personnels.

Personal information obtained in confidence

19 (1) Subject to subsection (2), the head of a government institution shall refuse to disclose any personal information requested under subsection 12(1) that was obtained in confidence from

- (a) the government of a foreign state or an institution thereof;
- (b) an international organization of states or an institution thereof;
- (c) the government of a province or an institution thereof;
- (d) a municipal or regional government established by or pursuant to an Act of the legislature of a province or an institution of such a government;
- (e) the council, as defined in the Westbank First Nation Self-Government Agreement given effect by the Westbank First Nation Self-Government Act;
- (f) the council of a participating First Nation as defined in subsection 2(1) of the First Nations Jurisdiction over Education in British Columbia Act; or
- (g) a First Nation Government or the Anishinabek Nation Government, as defined in section 2 of the Anishinabek Nation Governance Agreement Act, or an Anishinaabe Institution, within the meaning of section 1.1 of the Agreement, as defined in section 2 of that Act.

Renseignements personnels obtenus à titre confidentiel

19 (1) Sous réserve du paragraphe (2), le responsable d'une institution fédérale est tenu de refuser la communication des renseignements personnels demandés en vertu du paragraphe 12(1) qui ont été obtenus à titre confidentiel :

- a) des gouvernements des États étrangers ou de leurs organismes;
- b) des organisations internationales d'États ou de leurs organismes;
- e) des gouvernements provinciaux ou de leurs organismes;
- d) des administrations municipales ou régionales constituées en vertu de lois provinciales ou de leurs organismes;
- e) du conseil, au sens de l'Accord d'autonomie gouvernementale de la première nation de Westbank mis en vigueur par la Loi sur l'autonomie gouvernementale de la première nation de Westbank;
- f) du conseil de la première nation participante, au sens du paragraphe 2(1) de la Loi sur la compétence des premières nations en matière d'éducation en Colombie-Britannique;
- g) du gouvernement de la première nation ou du gouvernement de la Nation des Anishinabes, au sens de l'article 2 de la Loi sur l'accord en matière de gouvernance conclu avec la Nation des Anishinabes, ou d'une institution anishinabe, au sens de l'article 1.1 de l'accord, au sens de l'article 2 de cette loi.

- The test to determine when information is about a particular individual was stated by Justice Gibson in *Gordon v Canada (Health)*, 2008 FC 258 [*Gordon*], at paragraph 34:

“Information will be about an identifiable individual where there is a serious possibility that an individual could be identified through the use of that information, alone or in combination with other available information.”

(emphasis added)

- In *Public Safety*, Justice McHaffie defined a “serious possibility” as:

...a possibility that is greater than speculation or a ‘mere possibility,’ but does not need to reach the level of ‘more likely than not’ (i.e., need not be ‘probable on the balance of probabilities). Applying such a standard recognizes the importance of access to information by not exempting information from disclosure on the basis of mere speculative possibilities, while respecting the importance of privacy rights and the inherently prospective nature of the analysis by not requiring an unduly high degree of proof that personal information will be released (para 53).

[49] The parties do not dispute that this is the applicable legal framework. They disagree regarding its application to the facts of the case, to which we now turn.

[50] At this stage, a reminder is in order regarding the specific requests in issue, and what remains in dispute. The Hayes request seeks the addresses of all licensed personal production relating to relatively large quantities of marijuana (244 or more indoor plants, and/or 95 or more outdoor plants, and/or 35,625 grams in storage). The Cain request seeks a document in sortable format showing the FSAs of personal or designated producers, as well as similar information

about “registered users” – which Health Canada interpreted as meaning either a personal or designated producer.

[51] As to what remains in dispute, the Information Commissioner accepts that certain portions of the records in issue should not be disclosed because they contain personal information; this includes the full addresses, as well as some FSAs that refer to locales with small populations. Health Canada agreed to disclose the first character of the FSAs contained in the records. Therefore, neither of these points is in issue.

[52] However, the Information Commissioner does not accept Health Canada’s assertion that the second and third characters for other FSAs, as well as city names, should be protected from disclosure because there is a serious possibility that such data, when combined with other available information, could result in the identification of individuals.

[53] In addition, the Information Commissioner does not accept Health Canada’s claim that it would be unreasonable to ask it to review each FSA, to determine the risk associated with the release of the second and third characters. This point is discussed below, in connection with the second issue.

[54] The dispute at the centre of this case is about whether the second and third characters of FSAs with larger populations as well as city names, are protected from disclosure because there is a “serious possibility” that this data can be linked with other information to identify specific individuals. Related to this is the proper approach to assessing the risks regarding what are

referred to as “structured data sets” and the methodology to assess such risks. The parties and the intervener made submissions on this question, which I discuss below.

B. *Is the Minister authorized to refuse to disclose the records at issue pursuant to subsection 19(1) of the ATIA, because they constitute personal information?*

[55] The onus lies on Health Canada to establish that it was authorized to refuse disclosure of the records, and thus it is appropriate to begin with their position, even though they are technically the Respondent in this proceeding. This will be followed by a summary of the Applicants’ position and a discussion of the Intervener’s submissions.

(1) Health Canada’s Case

[56] Health Canada argues that it has met the *Gordon* test by demonstrating that there is a serious possibility that releasing the disputed information could result in the identification of individuals, because it could be combined with other available information. Its evidence included two affidavits and an expert report. These are described in some detail below, because they constitute the bulk of the evidence before the Court on this *de novo* review.

(a) *The Garrah Affidavit*

[57] Health Canada submitted an affidavit of Joanne Garrah, the Acting Director of the Licensing and Medical Access Directorate within the Controlled Substances and Cannabis Branch. She describes the evolution of the regulatory regime for medical cannabis, the type of

information that an individual must provide to obtain a license, as well as the information Health Canada publishes on its website on this subject.

[58] It is not necessary to review the history of the regulation of medical marijuana in detail; the relevant aspects of the regulatory regime in place at the time of the ATIP requests in this case have been described above.

[59] The affidavit sets out the information a person had to submit in order to obtain a registration for cannabis for their own medical purposes, including: their name, address, date of birth, and gender; their full address, telephone number, and email address; the full address of the site where the cannabis production would occur; whether it would be cultivated indoors or outdoors (and if the latter; that the location was not near a school, playground, or other place frequented by persons under 18 years of age); and whether the person would cultivate it themselves or designate someone else to do so. In addition, the person had to obtain a document from a medical practitioner specifying the amount of cannabis prescribed for daily use.

[60] The registration form included a privacy notice, which explained that: (i) the *Privacy Act* governs the use of the personal information that was being provided; (ii) the information “may be shared with law enforcement entities to confirm your lawful possession and production of cannabis...”; and (iii) “[i]n limited and specific situations, your personal information may be disclosed without your consent in accordance with subsection 8(2) of the *Privacy Act*”, and set out the person’s rights under that legislation.

[61] The affidavit also describes the types of information Health Canada collected and published on its website, including: monthly updates on the amount of cannabis sold for medical purposes; the number of personal or designated production registrations by province; the number of applications processed each month; and information for health care practitioners on medical cannabis.

(b) *The MacAndrew-Donnelly Affidavit*

[62] Health Canada also submitted an affidavit of Cassie MacAndrew-Donnelly, the team leader at the ATIP unit of Health Canada who has been involved in all ATIP requests regarding medical cannabis licenses since 2010. She has also been involved in all ATIP requests concerning FSAs and medical cannabis since Health Canada started receiving them in 2011.

[63] The affidavit reviews the processing history of the three ATIP requests underlying this case and then describes the other information that is available, which underpins Health Canada's rationale for refusing disclosure. The affidavit refers to three other sources of information: data that Health Canada publishes on its website concerning medical cannabis; other sources of publicly available data about populations including the 2016 Census Data and the Statistics Canada report "Population and Dwelling Count by FSA"; as well as other publicly available information including data released in response to numerous prior ATIP requests about medical cannabis registrations.

[64] The ATIP requests and releases described in the affidavit cover a period from April 2012 to January 2020 (i.e. subsequent to the dates of the Hayes/Cain requests, a point discussed

below). The affidavit indicates that a wide range of requests have been made, some of which seek national data while others focus on specific cities. Several early requests seek information by FSA, such as the April 2012 request, where the FSA for communities with a population of 6,200 or more was released. That cut-off was chosen because it corresponded to the average population of an FSA at that time. That release included the type of license, medical condition (with rare conditions removed), dosage, and the issue date of the license. A 2014 release provided the year of birth, dosage, sex, medical condition (rare conditions removed), and province (city removed) of individuals with a medical marijuana license. In November 2015, a release provided information for FSAs of communities with a population of 60,000 or more, and it also included medical condition (with rare conditions removed), dosage, province, type of license, and issue year.

[65] The affiant explains the concern about the “mosaic effect” created by such an accumulation of releases:

Health Canada ATIP is concerned that the current files under complaint, as well as the new ones that came in following these, show that requesters are trying to gather more and more bits of information over time in order to paint an even broader picture on the profile of medical cannabis licensees in Canada using the data points that have already been released. There has been a history of efforts to collect, link and publish information relating to licensees for medical cannabis. The link between all of these files could allow one to identify an individual if they were to release more than the first digit of the FSA.

[66] An example of this “history of efforts” is the interactive map of Canada that was made available on the internet, showing FSAs where medical marijuana applications were made

between 2001 and 2007. The affiant indicates that “(t)he map used a colour gradient which showed that the darker the area, the more medical marijuana licenses are granted there.” She goes on to describe the results of her exploration of the interactive features on the map. For example, when she clicked on a city name, a map of that city appeared divided by FSA, and when she clicked on a particular FSA, the number of medical marijuana licenses applied for in that location appeared, along with the patients’ medical condition(s) associated with that license. An Edmonton Journal article about the interactive map provided a link to a database that allowed users to search by medical condition, postal code, doctor’s specialty, daily dosage, and allowed storage of marijuana.

[67] The affidavit also explains the evolution in Health Canada’s approach to ATIP releases relating to medical marijuana licenses, and its growing concerns about the risk of re-identification:

At the time of the first medical cannabis ATIA request (A-2011-000945), Health Canada ATIP had established that the average FSA population in Canada was approximately 6200. It was decided that any community with less than 6200 would be protected as it represented a risk of identification, in conjunction with the other personal details being released. This request generated the interactive map described above.

By the time of the second request in 2014 (A-2014-000167) Health Canada ATIP was exercising their duty to assist and releasing as much information as possible without risking a privacy breach, while taking into account the first release under A-2011-00945. The information available in the public domain pertaining to the profiles of the licensees and the mosaic effect was not a risk at the time.

By the time of the third request in 2015 (A-2015-000332), Health Canada ATIP began to recognize that more and more information was beginning to exist in the public domain pertaining to medical

cannabis licensees. With the layering of information being requested, including licenses by FSA, and considering the other information that was already released such as gender, medical condition, age, and dosage, it became apparent that linkages could be made, especially if you happened to live in that FSA. A decision was made to protect the FSAs of communities under 60,000 in population (instead of the previous 6200) as the risk of identifying an individual within smaller communities in conjunction with the other disclosed information was too high.

It was also at this time that a requester in a 2015 request pointed Health Canada towards the ‘interactive map a colleague had created with the data received [through an earlier access request]’ as an example of the type of information they wanted. This was when Health Canada ATIP became very concerned with the amount of information being disclosed and what could be achieved with it over time.

[68] Part of Health Canada’s concern can be traced to the size of some FSAs:

The Statistics Canada 2016 Population and Dwelling Count by FSA Tables reveal that 25 geographical areas (as represented by the FSA) have populations of less than 100 people, while 20 geographical areas have a population of 25 people or less. For example, the FSA for geographical area E2R has [...] 5 private dwellings and a population of 10 people. The FSA for geographical area GIA has 1 private dwelling and a population of 1.

[69] In one particular FSA in [REDACTED], the number of registered users is the same as the number of personal producers, with no designated producers in that location. This means that anyone known to be producing cannabis with a licence is producing it for their own use for a medical condition. This location is home to a [REDACTED], which the affiant states increases the likelihood that a particular individual licensed to grow medical marijuana could be identified.

[70] Similarly, a previous ATIP release shows there is one person with a license for a daily limit over 100g – this person has a license for personal production and is authorized to possess large quantities. The release also indicates that the individual is male, born in 1984, and has severe arthritis. The affiant states that releasing the three characters of the FSA would narrow the specific locale where this person lives, increasing the likelihood that the person can be identified.

[71] Finally, the affidavit describes Health Canada’s rationale for not severing some of the information. This is described in more detail in the second part of the decision.

(c) *The Expert’s Report*

[72] Health Canada retained an expert, Dr. Khaled El Emam, who filed a report entitled “Privacy Risk Assessment for Data Releases about Registered Users and Producers of Cannabis in Canada.” The Executive Summary explains that the purpose of the report is to:

...define a framework for evaluating the re-identification risk in data about registered users and producers of cannabis. This framework defines the scope of what risks should be [assessed and] the necessary assumptions, as well as a methodology that should be followed to assess these risks.

[73] The expert states that the “actual risk of re-identification is a function of the assumptions that one is willing to make about possible adversaries [meaning those who might seek to use the data] and their knowledge...” The report considers two assumptions: a “permissive” one in which the adversary does not know who is in the dataset, and a “conservative” one in which the adversary knows who is in the dataset or that a specific individual is in the dataset. After

explaining some foundational elements of how such risk assessments work, the expert then applies his methodology to the two Cain releases to show the risks associated with the release of the three characters in the FSA. This part of the report is discussed in the Analysis section below.

[74] At the outset, two points from the Report should be emphasized. First, the expert uses the term “adversary” in a somewhat unusual way; it does not refer to an opponent or enemy (as the term is generally understood), but rather simply refers to someone who may seek to use the data that is released, whatever their motivation. The expert explains that “adversary” is the term generally used in the literature on this subject, and he uses it to avoid the potential confusion of introducing new terminology (see, for example: Information and Privacy Commissioner of Ontario, *De-identification Guidelines for Structured Data*, June 2016, at p. 2). I will use it in the same manner and for the same purpose in these reasons.

[75] Second, the expert underlines that any information from the datasets that is released will effectively become public because no further controls can be imposed on it once it is released. The clearest example of this is the interactive map described earlier: once the dataset underlying that map was released, individuals were free to share it, link it, and to make it public, and there was no practical way for Health Canada to stop or limit this.

[76] Turning to the background elements set out in the report, the expert offers a description of some foundational concepts that are used in the model he employs. A brief discussion of some of these will be helpful to an understanding of the parties’ submissions and the analysis that follows.

[77] The starting point for the expert is that “(w)hen data is released pursuant to an access to information request, the appropriate risk model to use is maximum risk... Following this model, we would estimate the risk for each record in a dataset and then assign the highest risk value of a record to the whole dataset. Therefore, the dataset risk is equal to the highest-risk record in the dataset.” The reason for this is that once the dataset is released, no further controls can be put on the use of the data to manage any associated privacy risks.

[78] The expert’s methodology relies on several key concepts, which are summarized below:

- *Represented population* refers to the subset of the total population that can realistically be in the dataset; in this case, the starting point is the population of each FSA within the datasets. The size of this group also depends on whether we assume the adversary knows someone who is in the dataset or not;
- *Quasi-identifiers* are the variables in the dataset that the adversary might know, e.g. age, sex, medical condition, postal code, or FSA;
- *Learning something new* – the risk of disclosure only pertains to information that would add to the adversary’s existing knowledge; if the relevant information is already known to the adversary, even though it may technically be categorized as personal information, the risk of releasing that particular data is not meaningful;

- *Equivalence class* refers to the size of the group with the same values on the quasi-identifiers. If the adversary knows who is in the dataset, the equivalence class size can be computed from within the dataset. If the adversary does not know who is in the dataset, then the equivalence class size must be computed from the relevant population: e.g. the number of males living in a particular FSA. The probability of disclosure is assessed by combining all of the relevant quasi-identifiers, because that is how an adversary is likely to use the data; and
- *Threshold for Identity disclosure* refers to “(t)he threshold that can be used to evaluate whether the risk of identity disclosure is acceptably small [and] is based on the size of the equivalence class.” For public data releases, the typical threshold used for determining if the group size is too small is 11, so that if a group count is lower than 11, the risk of identification is deemed to be too high. Other thresholds can be used, but 11 is the standard used by Health Canada for public release of clinical trial data.

[79] The final point worth noting here is that the expert’s methodology proceeded on the assumption that, in light of the data that had been released in previous ATIP requests, the relevant quasi-identifiers for the releases in issue in this case are year of birth or age range and sex, plus FSA and city.

[80] The expert's report then applies these concepts to the dataset for the Cain request, and concludes that whereas there are a number of FSAs that are high risk if either three or two characters of the FSA are released, there are no high risk FSAs if only the first character is released. The other parties question some of the expert's conclusions, and so this part of the report will be discussed in more detail below. This is sufficient to set the stage for a review of Health Canada's arguments.

(d) *Health Canada's Submissions*

[81] Health Canada argues that it has met the *Gordon* test because the evidence shows there is a serious possibility of identification when more characters from the FSA and the names of the relevant cities are combined with other available information. It submits that focusing solely on population size, as the Information Commissioner urges, is overly simplistic and ignores the real risks that exist due to the availability of other information and the obvious motivation of some requesters to combine this data.

[82] The Respondent's position is based on the idea of the "mosaic effect" – that data can be combined to reveal more than each single piece considered alone might show. It points to previously released data relating to medical cannabis licenses, including medical conditions, dosage, type and issue date of the license, year of birth, and gender of the licensed individual. Health Canada also relies on other ATIP requests that had been received subsequent to the Cain/Hayes requests as a further indication of the expanding nature of the elements of information that could be combined to identify someone.

[83] As for motivation, Health Canada contends that the pattern of requests combined with the creation of the interactive map demonstrates a concerted effort to collect, link, and publish information relating to medical cannabis licenses. It underlines that users could search the site connected to the map by medical condition, postal code, doctor's specialty, daily dosage, or allowable quantities for storage.

[84] Regarding the second and third characters of FSAs, Health Canada points to the publicly available Statistics Canada report that shows some FSAs have tiny populations, in communities where residents likely know each other. The Respondent cites one particular example from the evidence: one FSA has the same number of personal producers and registered users, and no designated producer. Health Canada argues that releasing further data on this FSA would allow anyone in that area who knows that someone has a license to confirm that this person is growing marijuana for their own personal use to treat a medical condition. It could also be linked to other available data to reveal other personal information.

[85] Health Canada also points to the visual and olfactory (smell) factors associated with cannabis production, which can be combined with other evidence to identify particular individuals. It points to a newspaper article regarding complaints from neighbours about the smell emanating from a house where marijuana was allegedly being grown under a medical license.

[86] Based on all of this, Health Canada asserts that it cannot safely rely on population size alone as a means of dealing with the risk of identification of licensed individuals. In this regard,

the facts of this case are similar to *Gordon*, where the Court found that combining the public information with the further information in the released data created a serious possibility of revealing personal information.

[87] Health Canada relies on the findings in the expert report, in particular the risks that arise under the “conservative assumption” that a person seeking to identify someone knows that the person is in the dataset. The analysis in the expert’s report shows that the re-identification risk increases as more FSA characters are disclosed, and it further increases if age and gender are included. Given the highly personal nature of the information in question, and its obligation to protect that data, Health Canada argues that its refusal to release more information is justified.

(2) The Applicants’ Submissions

[88] The Information Commissioner argues that Health Canada has failed to meet its burden of demonstrating that the refusal to release more information was authorized by subsection 19(1) of the *ATIA*. She submits that the evidence points to a “mere” possibility, not a serious possibility of identification.

[89] The Information Commissioner asserts that population is the key variable, as demonstrated by Health Canada’s previous use of population cut-offs to determine whether to release records. She argues that the use of this factor has been endorsed by the Supreme Court of Canada in a case where the Ontario Information and Privacy Commissioner ordered the disclosure of the number of registered sex offenders for all Ontario FSAs: *Ontario (Community*

Safety and Correctional Services) v Ontario (Information and Privacy Commissioner), 2014 SCC 31 [*Community Safety*].

[90] The Commissioner submits that the evidence relied on by Health Canada is speculative and insufficient to demonstrate a serious possibility of re-identification. The only evidence about the nine prior *ATIA* releases relied on as core elements of the information that could be linked is the general description in the MacAndrew-Donnelly affidavit. While this comes closer to meeting the test, the Information Commissioner argues it nevertheless falls short. The affidavit provides only four pages from one release made in 2014, but this one example is an outlier given the size of the FSA in question. Without access to the information disclosed in response to these previous requests – copies of which are in Health Canada’s possession – the Applicants contend that the record is inadequate.

[91] On the expert’s report, the Information Commissioner makes two key observations: first, Health Canada did not ask the expert to actually conduct any specific linking to demonstrate what was possible, and second, the report shows that the risks of re-identification are not serious for many of the FSAs. Further problems arise using the expert’s methodology because the datasets produced in response to the Cain request are incomplete (they do not include individuals licensed under a prior scheme, the *Medical Marijuana Access Regulations*, SOR/2001-277). In addition, the datasets are outdated, and the Information Commissioner points out that the evidence shows that many more people are now able to grow marijuana for personal use. Therefore, the risk of releasing the information is diminished because it would be even more difficult to link any particular location to a prior license for medical marijuana.

[92] The Applicants also assert that Health Canada cannot rely on statements about information disclosed on its website because the website pages are not in evidence, nor can it rely on visual or olfactory indicators because it has failed to take into account the number of illegal grow operations in Canada. In addition, the sight or smell at a particular place may not link the cultivation to a specific person because individuals can be licensed to grow at another location and can designate someone else to do it for them. It may also be an illegal grow-operation, although Health Canada has failed to produce any evidence about the prevalence of this. Similarly, the Information Commissioner challenges Health Canada's reliance on subsequent *ATIA* requests, because the records relating to these were not already in the public domain. If Health Canada can demonstrate a serious possibility of re-identification in relation to one or more of these later releases, it could justify a refusal at some point in the future.

[93] The bottom line for the Information Commissioner is that Health Canada was required to provide evidence showing actual linkages, and its failure to do that means that it has not met its burden of establishing a serious possibility of re-identification.

(3) The Intervener's Position

[94] The Privacy Commissioner did not take a position on the facts, but rather focused his submissions on the factors for assessing whether releasing the full FSA meets the serious possibility test, and if so, to what extent section 25 of the *ATIA* required redactions or other techniques to allow disclosure of more information (discussed in the next part).

[95] While underlining the importance of the broad definition of personal information that has been confirmed by the jurisprudence, the Privacy Commissioner acknowledges that “(a)dvancements in technology combined with the proliferation of public or quasi-public data sources magnify the potential for re-identification of datasets unless sufficient precautions are taken” (Intervener’s Factum, para 14).

[96] The Privacy Commissioner contends that the specific issue is novel, and prior cases are of limited assistance. In particular, the Supreme Court of Canada’s decision in *Community Safety* should be approached with caution because the underlying facts and evidentiary record were materially different from the present proceeding. In particular, in that case the evidence related to risk of re-identification was limited to newspaper articles and generic scholarly research, and there is no indication that expert evidence about equivalence classes or information about previous ATIP releases was before the Court.

[97] Turning to the expert report, the Privacy Commissioner supports the use of equivalence classes as a measure of the risk of re-identification for the type of dataset in this case: “Among other things, it provides an objective, transparent and logical framework for analysis based on the concept of risk and accords with generally accepted practices for assessing re-identifications risks” (Intervener’s Factum, para 21). While an equivalence class analysis may not be required in every case, the Privacy Commissioner asserts, “it is particularly useful for structured datasets dealing with sensitive information, especially where there may be a motivation to re-identify individuals” (Intervener’s Factum, para 22). While the risk of re-identification will never be eliminated, the Privacy Commissioner’s position is that the equivalence class analysis combined

with a “motivated intruder” test is an appropriate methodology for assessing the magnitude of the risk to determine whether the serious possibility test has been met.

[98] The Privacy Commissioner also supported the use of other “quasi-identifiers” as relevant variables in assessing the risks associated with the mosaic effect. The number of quasi-identifiers depends on both the type of available information and its relevance to the type of mosaic that is of concern. The basic idea is that the more information the adversary has, the greater the risk of re-identification. Therefore, an increase in the number of variables to be considered leads to a corresponding rise in the chances that an unacceptably small combination of values among them can result in the re-identification of a specific individual.

[99] The Privacy Commissioner submits that the persuasiveness of this evidence depends on the extent to which the relevant data can be linked. Thus, for example, if the relevant population has changed significantly over time, or if the datasets include different, non-comparable variables, the possibility of linkage by connecting earlier data to more recent information is diminished.

[100] Concerning the question of “response knowledge” (i.e. whether the adversary knows that an individual is in the dataset), the Privacy Commissioner submits that this assumption only holds if there is “a plausible scenario that would allow for one or more adversaries to infer that a person is in the dataset” (Intervener’s Factum, para 38). This could include a “nosy neighbour” relying on personal observation and other available information. The analysis should also take into account multiple adversaries because once the data is public there is no restriction on its

further distribution. However, the plausibility of such an assumption declines if there are too many confounding variables such that it would not be reasonable to infer that identification is reasonably possible.

IV. Discussion

[101] Although the specific issue raised by this case has not been addressed before, the general principles that orient the analysis are well established and worth repeating.

[102] First, access to information is a foundational right, essential to the health of our democracy and to improving the quality of public administration in Canada, which in turn sustains the public's trust in government institutions. There are many legitimate reasons for inquiring about how the administration of the medical marijuana licensing regime operates, and greater public awareness of such matters must be presumed to be a good thing.

[103] Second, the protection of privacy is also a hallowed value; the right to privacy is given constitutional and legislative protection because of its importance to individual dignity. The protection of personal information is an essential element of individual dignity, and core to that is the right of individuals to choose whether, when, and how to share information about themselves with others. There are many legitimate reasons why a person may not wish others to know that they are using medical marijuana because of a medical condition, and Health Canada was rightly concerned about protecting the information it held about such matters.

[104] Third, the “serious possibility” test set out in *Gordon* is still the governing authority that all parties submit should guide the analysis. This is because *Gordon* recognizes that information that is not inherently personal may be combined with other available data to create a serious risk, whereby the mosaic created by such efforts could lead to the identification of a specific individual. I agree that this is the applicable test here.

[105] Applying these principles to the facts of this case brings us to the crux of the matter. Has Health Canada demonstrated that disclosing the second and third characters of the relevant FSAs, and/or the names of the cities, creates a serious possibility of re-identification?

[106] A number of considerations lead me to conclude that Health Canada has met its onus.

[107] First, it seems to me that the type of personal information in question is a central concern for this type of analysis. Government agencies hold all sorts of information about individuals, and while all information that qualifies as “personal” under the statutory definition merits protection, it must be acknowledged that the disclosure of some particularly sensitive types of personal information can be expected to have particularly devastating consequences. Information about an individual’s medical condition(s) must rank very high on any such list: it is among the most intimate information any of us possess, and the decision of whether or when to share it, and how much to disclose, can be a gut-wrenching choice, with significant consequences for the individual, their family, and friends.

[108] Flowing from this, the risks of disclosure of such intimate information must be reduced as much as is feasible. This is not to suggest that Health Canada or any other government entity can guarantee that such information will never be disclosed; the law does not seek a type of certainty that can never realistically be obtained. However, this approach supports Health Canada's assertion that it took an appropriately restrained approach to disclosure, and it is pertinent to a consideration of the expert's report, discussed below. On this point, it should be emphasized that Health Canada was under an obligation to try to prevent the disclosure of every individual's personal information. Even though it held thousands of records, the obligation was towards each individual.

[109] Next, I find that the existence of the interactive map, and Health Canada's legitimate concerns about what had been done with the information released in prior releases are important considerations in assessing whether disclosing the information requested by the Applicants created a serious risk of re-identification. It may not always be essential to provide concrete examples of the motivation that some individuals may have to connect information or the feasibility of such efforts. However, the existence of evidence demonstrating that connections among disparate pieces of relevant information have previously been made and that the results have been made available to the public is a relevant consideration in applying the serious possibility test. The information previously released must be assumed to still be available, even if the website is no longer accessible online.

[110] Although I accept that Health Canada's reliance on past disclosures was legitimate and appropriate, I am not persuaded that Health Canada's reference to the subsequent ATIP requests

was a legitimate consideration, because there is no evidence about what, if any, further information was disclosed pursuant to those requests during the intervening period. It is worth unpacking this point.

[111] As noted earlier, the MacAndrew-Donnelly affidavit summarized information that had already been disclosed at the time the Cain/Hayes requests were made. In my view, this was a relevant and appropriate consideration in assessing the mosaic effect, because that information was already in the public domain.

[112] However, the affidavit also refers to subsequent requests made after the Cain/Hayes requests, and the record is not clear whether any other information was disclosed in response to these requests, either prior to Health Canada's final refusal, or between then and the time of the hearing of this matter. Absent that evidence, I agree with the submissions of both the Information and Privacy Commissioners that the impact of any future disclosure was merely hypothetical at the time of Health Canada's refusal, and it remains so for the purposes of my *de novo* review. The fact that a more complete mosaic may be created by future releases is both true and irrelevant, because Health Canada has an ongoing obligation to assess the risks, and if at some future point it concludes that the accumulation of information released created a serious risk, it could refuse to disclose the information that tipped the balance. A future potential risk that can be mitigated by decisions that are within Health Canada's control cannot be invoked to justify the refusal in issue here.

[113] Turning to Health Canada's evidence, while I agree in principle with the Information Commissioner's submission that Health Canada could and should have filed more information to demonstrate the contents of the prior releases, I am persuaded that the details set out in the MacAndrew-Donnelly affidavit (leaving aside the references to the subsequent releases) are sufficient to demonstrate the type of information that could form the building blocks of the mosaic of information about specific individuals licensed to produce medical marijuana. On this point, it is important to note that the affiant was personally involved in all of the prior ATIP requests, and so was able to describe the information in the records that were released in great detail.

[114] The affidavit shows the progressive release of more information about medical marijuana licenses, as well as details about the individuals who received them, including medical conditions, year of birth, gender, type of license issued, and dosage. Two main issues were raised about this information.

[115] First, the Information Commissioner questioned whether the data was comparable across the various releases, because if there are substantial differences between the various datasets, the risk of an accurate linkage would be reduced or eliminated. Related to this is the question of the completeness of the data, because it is acknowledged that some of the relevant data sought in the Cain requests was contained in a separate database and was therefore not included. The key point here is that the prior database contained much more data than the one used to respond to the Cain request.

[116] Second, the Information Commissioner argued that the data assembled in response to the two requests was no longer accurate given the passage of time, and in particular the significant increase in the number of Canadians who can grow marijuana after legalization. It argued that on the *de novo* review, the Court was required to take into account developments subsequent to the time of the original Health Canada refusals, and this included the evidence that many more people are now growing marijuana for personal use. This diminished the chance that knowledge that someone was cultivating marijuana would create a significant risk of identifying licensed medical producers.

[117] Two main questions emerge from the submissions on this point. First, is the data comparable? While it would have been preferable to have evidence, from either one or more of Health Canada's affiants or the expert, in my view a reasonable inference can be drawn from the evidence that is in the record. The records all relate to the medical marijuana licensing regime, and the evidence is clear that Health Canada has required individuals to provide medical evidence of a health condition to justify their need to use medical marijuana under all of the different regulatory regimes. It is reasonable to infer from this that medical conditions that warranted a license under a prior regime would likely justify its continuance under the next one, and so on. It is also reasonable to infer that the vast majority of individuals who were motivated to seek a license under a previous regime would likely be interested in continuing it, to the extent new requirements were put in place under subsequent regimes. This is particularly the case because the Hayes request sought information relating to large quantities, which is a sub-set of the larger group of licensed individuals and designated producers. An individual would need medical evidence to justify a license to grow such a quantity of medical marijuana, which in turn

makes it more likely that they would seek to continue to obtain the necessary authorization. I am satisfied that in the particular circumstances of this case, taking a practical view of the matter, the evidence shows that the datasets are likely broadly comparable.

[118] On this point, it is important to add that even if the datasets are not entirely comparable, it is reasonable to infer that there is a significant degree of continuity in the licensed population that is included in them. The key point is not that the data is statistically comparable for the purposes of scientific or social science research. Rather, the question is whether there is a significant possibility that this data can be combined to identify particular individuals. From the privacy perspective, I am persuaded that the datasets are sufficiently comparable to serve as a foundation for assessing the risk that a mosaic of information could be assembled.

[119] I am also satisfied that even though the dataset produced in response to the Cain request was not complete, the evidence does not demonstrate that the absence of the data about users licensed under the prior regime would significantly reduce the risk of re-identification. For the reasons set out in the previous paragraphs, it is likely that many, if not most, of those licensed under the previous regulations are included in the *ACMPR* data that is relevant here. I also note that during the course of the investigation, the Information Commissioner accepted Health Canada's explanation for not including the previous data in its response to the Cain request. Once again, the fact that the datasets may not be exactly comparable might be a problem for a statistician or social scientist, but it is not an impediment to a motivated user seeking to identify a person who was licensed for personal production or a designated producer under the medical marijuana licensing regime.

[120] Second, what is the relevant timeframe for the Court conducting the *de novo* review? Am I to assess the information that was public at the time of Health Canada's refusal, or should subsequent events be taken into account? This is particularly salient in this case, because of the time that has passed since the original requests and because in the interim period the legalization of marijuana has changed the underlying factual matrix. The access requests were made in August and October 2017; Health Canada's final refusal to disclose was dated January 20, 2020; and the case was heard in February 2022. During that time-frame the number of individuals who cultivated small amounts of cannabis for personal use continued to increase, as did the number of individuals registered with Health Canada for personal and designated cultivation (as of June 2020, the latter category included 33,614 individuals). Both facts are pertinent to the assessment that the risk of disclosure today would pose, but they may not have been representative of the risk as of the date of Health Canada's refusal.

[121] In my view, in undertaking the *de novo* review mandated by the *ATIA*, I am required to take into account more recent developments insofar as they are pertinent to the task at hand; namely, assessing whether the disclosure of further information from the records creates a serious possibility of disclosure of personal information about a specific individual. In *Concord*, I found that the passage of time was a relevant consideration in assessing the risk of a disclosure, albeit in a completely different context (at paras 82-85). I remain of that view.

[122] In the same way that Health Canada was required to take into account any relevant developments between the date the access requests were filed and the time of its final decision (e.g. whether any other information had become public), I find that a court conducting a *de novo*

review of a refusal to disclose a record should take into account any relevant changes between the date of the refusal and the time of the hearing of the matter. Failure to do so would artificially freeze time for no purpose, and would be inconsistent with the *de novo* nature of the independent review a court is mandated to undertake. The assessment must be a practical one, taking into account all relevant evidence as of the date of the court hearing.

[123] For example, if Health Canada produced evidence at the hearing of information releases, or efforts to link available information that occurred subsequent to its final decision on the Cain and Hayes requests, this would be relevant to the risk assessment I am required to undertake in the context of the *de novo* review. The same is true about the evidence of the increasing prevalence of marijuana cultivation under the legalization regime.

[124] On this final point, a caveat is in order. The passage of time does not erase the possibility that the datasets that were previously released could be combined with the complete FSAs and city names requested in this case, to create a significant risk that individuals could be identified. Absent evidence to the contrary, the presumption must be that the information that was previously disclosed is still available and can be combined with other, more recent information. In this regard, the increase in marijuana cultivation is a relevant consideration, but it is not determinative.

[125] Next, I am not persuaded by the Information Commissioner's assertion that Health Canada's previous reliance on population thresholds indicates that it should apply this criterion

to the releases in this case. I also do not find that the *Community Safety* decision is particularly persuasive on this point.

[126] First, the MacAndrew-Donnelly affidavit as well as Health Canada's final refusal letters explain why they concluded that population thresholds were no longer sufficient to manage the risks associated with these releases, and I find their rationale compelling. There is no doubt that the obligation to assess privacy risks is ongoing, and that Health Canada was required to consider developments that occurred subsequent to its prior releases. The pattern of requests, their similarity and specificity, and the emergence of the interactive map are all factors that supported Health Canada's conclusion that population thresholds alone were no longer sufficient to protect the sensitive personal information contained in the datasets. I find no fault with this approach.

[127] Second, I am not persuaded that the *Community Safety* decision supports the use of population thresholds in this case. Each case must be assessed on its own facts and in light of the overall circumstances. While the population size of FSAs was an important consideration in *Community Safety*, several elements limit its application here. In that case, the record in question was found not to be personal information (see paras 17, 35, 36 and 64) and this finding was not challenged when the original decision was appealed (para 22). In addition, there was no evidence about how the record in dispute could be cross-referenced with other information in order to identify an individual (para 60), nor was there a pattern of multiple requests (para 62). Instead, all that was before the Court was "unconvincing and generic scholarly research on 'identifiability'" which did not address the specific facts of the case (para 60).

[128] For these reasons, I am not persuaded that this decision supports a more general proposition that population thresholds are suitable to manage privacy risks. Rather, it stands for a much more limited proposition that rests on the factual matrix before the Court, which is significantly different from the evidence in the record here.

[129] Instead, I find the facts in *Gordon* to be more similar to the case at bar. That case involved a challenge to Health Canada's refusal to release the province relating to reports of adverse drug reactions that was held in a database it maintained. Drug manufacturers were required to provide such information, and it was supplemented by reports from health professionals and consumers, who provided the information on a voluntary basis. The database maintained by Health Canada contained approximately 100 active data fields, of which 82 had been disclosed. However, Health Canada refused to release information about the province where the report was received (which was not necessarily the province where the adverse drug reaction actually occurred).

[130] Health Canada justified its refusal on the basis that the province was "personal information" because of the risk that it could lead to the identification of particular individuals if it was linked to previously disclosed information. The Information Commissioner agreed with Health Canada on this point.

[131] As noted previously, the Court set out the "serious possibility" test in *Gordon*, and applying that test to the facts before it, the Court upheld Health Canada's refusal to release the information. Two elements underpin the Court's decision: first, the fact that some provinces and

territories have relatively smaller populations, and second, the specific example of a case in which a Canadian Broadcasting Corporation reporter contacted a family to inquire whether their daughter's death was connected to an adverse drug reaction, based on information it had obtained from the database of adverse drug reactions as well as the daughter's obituary. Based on Health Canada's evidence, the Court concluded at paragraph 43 that disclosure of the province would:

substantially increase the possibility that information about an identifiable individual... would fall into the hands of person seeking to use the totality of the information disclosed from... the database, in conjunction with other publicly available information, to identify "particular" individuals.

[132] It is noteworthy that the Court was not persuaded that the lack of expert evidence undermined the privacy claims, nor did it accept the challenges to the data quality advanced by the applicants in that case, which included that there was significant under-reporting of adverse drug reactions, the database included suspected adverse drug reactions rather than only scientifically established instances, and there was a delay in reporting. The Court found that Health Canada had met its burden of establishing that it was authorized to refuse to disclose the record.

[133] This brings me to the expert's report. Three observations are in order. First, I am not persuaded that Health Canada's failure to ask the expert to conduct an actual linkage analysis using the available data is fatal to its case. Such evidence would undoubtedly have been useful in conducting the *de novo* review, but it is not mandatory, and its absence does not diminish the weight accorded to the expert's opinion. That opinion, combined with the reasonable inferences

about the continuity of the dataset, the example of the interactive map, and the highly sensitive nature of the information combine to overcome this deficiency. However, in future cases the failure to engage in such an exercise might well tip the balance in favour of disclosure.

[134] Second, the fact that the expert did not conduct an analysis on all three of the datasets produced in response to the two access requests is also not a fatal flaw. To the extent that the expert's opinion demonstrates a risk associated with further disclosure, it is sufficient to support a conclusion that applies across the datasets, given their similarity and the fact that the same concern about re-identification arises in relation to all of the records. If anything, the fact that the Hayes dataset is limited to licenses for large quantities – a sub-set of the wider database – amplifies the concern about re-identification.

[135] Third, while I accept and rely on many of the Privacy Commissioner's submissions, I do not accept that the evidence in this case is sufficient to draw any more general conclusions or to establish any general rules, for example regarding the appropriate size of an equivalence class. Such findings should either be the product of regulatory consultations or expert evidence on this specific point, and I am not prepared to make any general pronouncements based on the evidence before me. The expert's report addressed the situation in this particular case, and that is how I have treated it.

[136] That said, I am persuaded that the expert's report is both highly relevant and persuasive evidence regarding the risks associated with further disclosure of the second and third characters of the FSAs, and, by extension, the names of the cities. There is no doubt that the expert

possesses highly relevant expertise, and that his explanation and analysis was both thorough and compelling. Indeed, none of the parties took issue with the expert’s recommended analytical approach – for example, the Information Commissioner sought to rely on it to bolster their position. The main dispute between the parties concerns the appropriate assumptions to make, the risk tolerance to be applied, and the outcome of the analysis. I find that the analytical approach set out by the expert is suitable for the analysis of the risk of disclosure for the types of structured datasets involved in this case.

[137] As explained below, I accept that in assessing the risks in this case it is appropriate to take the “conservative assumption” the expert recommended, and I find that his analysis of the risks of disclosure is highly persuasive.

[138] Acknowledging that assessing the risk of re-identification associated with the release of a particular record involves a degree of uncertainty, it must be noted that such exercises are not unknown in law, in particular in administering access and privacy laws. The approach to applying the “serious possibility” test endorsed in *Gordon*, recently confirmed in *Public Safety*, provides the framework for addressing the uncertainty associated with the predictive element of the exercise. Justice McHaffie described the approach in *Public Safety*, at paragraphs 53-54, finding that the “serious possibility” test in *Gordon* means:

[A] possibility that is greater than speculation or a “mere possibility,” but does not need to reach the level of “more likely than not” (*i.e.*, need not be “probable” on a balance of probabilities). Applying such a standard recognizes the importance of access to information by not exempting information from disclosure on the basis of mere speculative possibilities, while

respecting the importance of privacy rights and the inherently prospective nature of the analysis by not requiring an unduly high degree of proof that personal information will be released.

Beyond this, it seems unnecessary, and may even be impossible, to try to further subdivide or parse the requisite degree of likelihood that an individual could be identified.

[139] It is obvious that the underlying assumptions that are used constitute an essential determinant in the assessment of risk, and these are a key focus of the parties. For example, a key assumption that is disputed among the parties is whether the adversary knows who is in the dataset; as discussed below, this is a key variable in the expert's risk assessment, and has an important impact on the outcome of the case.

[140] The expert's report does not suggest a preference regarding this point, but rather simply states:

Whether an adversary knows that someone is in the dataset is an important assumption that must be made in a re-identification risk analysis because it affects how the risk is calculated. There are reasons for each of these assumptions to be reasonable ones, and therefore we will make both assumptions and perform the analysis twice, once under each assumption (Expert's Report, p 7, p. 971, R. Record).

[141] The Information Commissioner argues that the expert's opinion rests on the assumption that a recipient of the disclosed data knows all or most of the individuals in the dataset, and submits that this is not a tenable approach. The Information Commissioner asserts that, at best, a recipient may know some of the individuals in the dataset. Further, the Information Commissioner says that under the expert's assumption, there are 46 instances in which Health Canada's disclosure of the first character of an FSA has created a high risk for the Cain

request, and a further 25 instances in relation to the Hayes request. The Information Commissioner describes these as “false positives” and submits that this further weakens the force of this assumption.

[142] I am not persuaded by this argument. I agree with Health Canada and the Privacy Commissioner that the expert’s analysis does not rest on the faulty proposition that the adversary “knows most of the individuals in the data set.” A careful reading of the expert’s report does not support that conclusion. Although there are references in the report to an assumption that “an adversary would know who is in the dataset” (Expert Report, p. 17), it is not reasonable to interpret this as meaning the recipient of the disclosure knows everyone in the dataset. In the earlier portions of the report it is clear that the relevant assumption is that the adversary either knows someone who is in the dataset because they have a license or are a designated producer, or the adversary knows some things that could be useful in seeking to identify a person who may be in the dataset (e.g. what FSA they live in, plus their gender, age, or medical condition, and/or that they use or cultivate medical marijuana). None of the examples used by the expert in the discussion refer to an adversary knowing everyone in the dataset, and therefore I do not accept the Information Commissioner’s argument on this point.

[143] As noted earlier, the expert does not provide an opinion on which of the two assumptions should be applied in this case, because he finds that “(t)here are reasons for each of these assumptions to be reasonable ones...” (Expert Report, p. 7). In the circumstances of this case, I find that it is more appropriate to assume that an adversary knows that a person is (or might be) in the relevant datasets.

[144] Several factors support this conclusion. First, given the highly sensitive nature of the information, it is appropriate to seek to lower the risk of re-identification to the maximum extent that is reasonably feasible, and the assumption of response knowledge helps achieve this. Second, unlike the *Public Safety* case, the types of other information that can be used to make linkages are not all confidential; for example, a person's general location (or specific address), as well as their gender and age range are the types of things that neighbours, friends, or family members may readily know. It is also true that such people may have been told by the person (or might suspect) that they consume marijuana, but not know that it is for a medical reason. Third, the pattern of requests and the existence of the interactive map show a certain motivation to glean more information about the administration of the licensing regime.

[145] The expert offers several examples that illustrate the concern, which he describes (using the lexicon of this type of analysis) as the "direction of attack":

An adversary can attack a dataset in two general ways. The first is if the adversary knows someone in the [represented] population, and then tries to find a matching record in the dataset. That target someone can be an acquaintance of the adversary. For example, the adversary may be trying to re-identify the record of a neighbour or a co-worker. Alternatively, the adversary may be trying to re-identify the record of a famous person (e.g., a politician or a sports personality).

The second direction of attack is when the adversary tries to match the records in the data with real people in the population. This is typically done by creating a population registry of some sort. The adversary can do that by using publicly available information, public registries, or social media, for example.

[146] In light of these considerations, I find it more appropriate to apply the assumption that best protects individual privacy, which in this case involves the adversary having some relevant knowledge that an individual is (or is likely to be) in the dataset. This may also include some information about one or more of the quasi-identifiers. Whether this is a “nosy neighbour” (to borrow an expression often used in these types of cases), an interested journalist, or a person seeking to identify locations where large quantities of marijuana are likely to be found (for sinister reasons or more benign purposes), it is reasonable to assume that at least some of the recipients of the records in question would possess relevant knowledge or be motivated to obtain it.

[147] A second assumption that is challenged is whether the risk analysis should include quasi-identifiers. This turns on whether it is reasonable to assume that linkages can be made between the data. The expert noted that prior releases included information about the year of birth or age range and sex of licensed individuals, as well as the FSA or city where they lived. The expert labels these as “definitive quasi-identifiers.” The Information Commissioner argues that these should not be factored into the risk analysis because it is not possible to compare the datasets used in the earlier releases with the information in the records here.

[148] The major differences between the datasets include the time frames they cover and the number of registrations recorded under the different medical marijuana regulatory regimes. At this point, I simply re-affirm that I am not persuaded by this argument, for the reasons set out at paragraphs 117-119 above. Given the information previously released, it is reasonable to assume that an adversary would be able to link enough of the data from the previous releases to the

current records, and this would increase the risk of re-identification. The risk is higher for the sub-set of cases involving large quantities. Therefore age (or age ranges), gender, and FSA or city are relevant quasi-identifiers for the purpose of this analysis.

[149] Applying these assumptions to the data in the records leads to the conclusion that releasing more than the first character of the FSA creates an unacceptable risk that individuals' privacy might be violated. Using the terminology from *Gordon*, endorsed by *Public Safety*, I find that Health Canada has established a "serious possibility" of re-identification, through evidence that rises above speculation or "mere possibility", even if it falls short of showing that such a result is "more likely than not." That is all the law requires to justify applying the class-based exemption in subsection 19(1).

[150] I base this conclusion on the combined force of the MacAndrew-Donnelly affidavit's description of the prior releases and the interactive map, as well as the expert's report. There is no need to repeat the discussion of the affidavit, but certain details of the expert's report merit further discussion here.

[151] The expert analyzed the datasets using the two assumptions: that an adversary does not know someone in the dataset, or that an adversary possesses such knowledge. As explained above, I find that the second assumption is applicable here, and so I will not discuss the other assumption nor the expert's corresponding analysis. The expert found that year of birth or age range, FSA, or city and sex were "definitive quasi-identifiers" and I accept this based on the prior releases as described in the MacAndrew-Donnelly affidavit.

[152] The expert also identified 57 high risk FSAs simply because they “have a low count (less than 11) on any combination of age and gender.” (Expert’s Report, p. 16) I accept that this is a relevant criteria, based on the expert’s analysis and the fact that Health Canada will not release similar data in other circumstances where the count is lower than 11.

[153] Turning to the specific datasets, it is important to underline that this portion of the expert’s analysis was focused on the size of the relevant FSA, and did not apply the other quasi-identifiers. As he explains:

Note that [in] this analysis we are only looking at the geography and do not have age and gender in the dataset. If these are added the equivalence class sizes would be lower (and hence risk levels would be higher). Under such conditions there would be even more records that are high risk.

[154] As the expert’s report demonstrates, even without considering the quasi-identifiers of age and gender, releasing either the full three characters in the FSA or only the first two characters results in a significant number of high risk locations being identifiable: over 1,000 FSAs are problematic if the complete FSA is released, while 82 locations are high risk if the first two characters are released. In contrast, the release of only the first character of the FSA, only 3 or 4 locations are problematic.

[155] In practical terms, the expert’s analysis concluded that if the adversary knows that someone is in the dataset, 611 out of 673 designated producers would be at high risk of re-identification, and for personal producers there would be 1011 problematic FSAs with 4060 of 11,100 individuals being at high risk for one of the Cain requests, while 4183 out of 11,841

individuals would be at high risk for the other request. It bears repeating here that if age and gender were factored into the analysis, the risks would be higher than those identified by the expert.

[156] In comparison, releasing only the first character poses a much lower risk: 20 out of 673 designated producers would be at high risk, and 12 out of 11,100, or 14 out of 11,841 personal producers would be at high risk.

[157] This confirms that releasing more than the first character of an FSA creates a significantly greater risk of re-identification. While I accept the Information Commissioner's statement that the expert's report tends to indicate that there is some risk associated with releasing the first character of the FSA, it does not follow that this would justify releasing more information if doing so increases that risk. Neither logic nor common sense support such an approach.

[158] Based on this analysis, it is also reasonable to conclude that releasing the names of the cities, with or without more information about specific FSAs, would increase the risk of re-identification. The main focus of the submissions and evidence of the parties was on the FSA data, but logic suggests that the same risks would result from disclosing the names of cities, because these can also be linked to other available information, including the first character of the FSA (which has already been released) to narrow down the area of focus for a motivated adversary. The fact that larger cities can be sub-divided by FSA is also pertinent, as demonstrated by the example of the interactive map.

[159] For all of these reasons, I find that releasing the second or third character of the FSA, or the names of the cities, would create a serious possibility of re-identification, and this information therefore falls within the definition of personal information about an identifiable individual. It follows that the records assembled by Health Canada in response to the requests should not be disclosed in their current form.

A. *Did the Minister correctly refuse to further sever the records pursuant to section 25 of the ATIA?*

[160] Health Canada agreed to the Information Commissioner's request to release the first character of the FSA in relation to both requests, but refused to undertake further analysis on the basis that doing so would impose an undue burden that went beyond what section 25 of the *ATIA* requires. The question before the Court is whether this is the right approach under the statute, in the particular circumstances of this case.

[161] A few matters are not in dispute. All parties agree that some of the remaining information should be redacted, including the full FSA or names of cities with small populations. In addition, no party suggests that Health Canada can reasonably conduct the type of risk analysis that would be required to identify higher- and lower-risk FSAs or city names without relying on software. However, the parties take different positions on whether Health Canada is required to create and apply such software in order to meet its severance obligations under section 25.

(1) The Position of the Parties

[162] The Information Commissioner submits that section 25 requires the severance and disclosure of all records not meeting the serious possibility test. Severance is mandatory, not discretionary, and the burden is on Health Canada to show that it is authorized to refuse disclosure. Reasonable severance, as required by section 25, is severance without serious problems, as indicated by the French version of the provision: “le prélèvement de ces parties ne pose pas de problèmes sérieux” (emphasis added by Applicants).

[163] According to the Information Commissioner, Health Canada can apply a population threshold analysis to each FSA, as it has done for previous releases. While this would require an analysis of each FSA, it would not involve the development of specialized software or be overly difficult. The resulting records would be meaningful, and severing the records by population threshold would not involve “serious problems.” In the alternative, the Information Commissioner supports the Privacy Commissioner’s assertion that other de-identification techniques could be used to release more of the information.

[164] Health Canada argues that it conducted a reasonable severance exercise consistent with its legal obligations, and requiring it to do more would impose an undue burden. It contends that any further severance would require it to undertake an analysis of the risks associated with releasing more information for each FSA, which cannot be done manually because of the overall complexity of this work.

[165] As for population thresholds, as discussed earlier, Health Canada argues that its prior experience has demonstrated why using population size alone is no longer sufficient to provide

adequate privacy protection. It points out that privacy is a primary consideration in severing information pursuant to section 25 (*Attaran*, at para 25). Because using a population threshold would not take into account the other relevant quasi-identifiers or the number of people in a particular record set, Health Canada argues it is therefore no longer suitable as a means of managing the risk of re-identification.

[166] On the interpretation of section 25, Health Canada rejects the Information Commissioner's reliance on the French version of the provision. It argues that Parliament's intent is revealed in the multiple associations between "severance" and "reasonable" in the *ATIA*, where the translation of the term "reasonable" is "efforts raisonnables." Health Canada submits that this confirms that it is only required to undertake reasonable efforts. In the alternative, it contends that the expert's report indicates that further severance would cause serious problems.

[167] Health Canada contends that analyzing identification risks across multiple data releases is unduly burdensome, because it must be done separately for each FSA, and would need to be repeated for each request. In order to do this, Health Canada asserts that it "would have to develop expertise in data and privacy analytics, or develop software to ensure their consistent application and to perform the necessary calculations" (Respondent's Memorandum of Fact and Law, para 111).

[168] The Privacy Commissioner suggests that other de-identification techniques can be used as an alternative to achieve the goals of section 25, noting that the term "severance" is undefined and traditional approaches to redaction may be ineffective in the context of structured datasets.

The Privacy Commissioner asserts that using more sophisticated techniques for severance could better serve the requirement under section 25 to release as much information as “can reasonably be severed” while fulfilling the Minister’s duty to “respond to the request accurately and completely” as set out in section 4(2.1). The Privacy Commissioner states that the appropriate technique would depend on the nature of the structured dataset, the overall context, and the institution could assess this with the assistance of technical experts.

[169] In response, Health Canada asserts that it cannot manipulate the data as suggested by the Privacy Commissioner, because doing so would violate section 67.1(1) of the *ATIA*, and following the suggested approach would have the perverse effect of releasing less data than what they ultimately released. Health Canada points to the Treasury Board Secretariat’s Privacy Implementation Notice 2020-03, which does not require the type of manipulation of the datasets recommended by the Privacy Commissioner. This Notice indicates that such methods may be appropriate in certain circumstances, such as releasing information relating to government audits or program evaluations and statistical reports, but are not permitted to be applied to records containing personal information. Instead, the Notice requires that severance (also known as redaction) be applied to personal information.

B. *Discussion*

[170] The core question here is whether more effort was required by Health Canada to respect its obligations under section 25. Although there is some guidance to be gleaned from precedent, in particular *Merck Frosst*, it does not appear that this specific question has been addressed in any previous decisions.

[171] Section 25 of the *ATIA* states:

Severability

25 Notwithstanding any other provision of this Part, where a request is made to a government institution for access to a record that the head of the institution is authorized to refuse to disclose under this Part by reason of information or other material contained in the record, the head of the institution shall disclose any part of the record that does not contain, and can reasonably be severed from any part that contains, any such information or material.

Prélèvements

25 Le responsable d'une institution fédérale, dans les cas où il pourrait, vu la nature des renseignements contenus dans le document demandé, s'autoriser de la présente partie pour refuser la communication du document, est cependant tenu, nonobstant les autres dispositions de la présente partie, d'en communiquer les parties dépourvues des renseignements en cause, à condition que le prélèvement de ces parties ne pose pas de problèmes sérieux.

Safety of individuals

~~25 The head of a government institution may refuse to disclose any personal information requested under subsection 12(1) the disclosure of which could reasonably be expected to threaten the safety of individuals.~~

Sécurité des individus

~~25 Le responsable d'une institution fédérale peut refuser la communication des renseignements personnels demandés en vertu du paragraphe 12(1) dont la divulgation risquerait vraisemblablement de nuire à la sécurité des individus.~~

[172] In *Merck Frosst*, the Supreme Court of Canada confirmed that this provision imposes a mandatory obligation on government institutions, and the onus rests on the institution to justify why it cannot disclose part of a record through reasonable severance. This involves both a semantic and cost-benefit analysis.

[173] The semantic analysis “is concerned with whether what is left after excising exempted material has any meaning” (*Merck Frosst*, para 237). This is not in dispute here, as it is evident

that disclosing the second and/or third character of the FSA, plus the names of some cities, would provide meaningful information.

[174] The focus of the arguments here is on the cost-benefit aspect, which is described in *Merck Frosst* at para 237:

The cost-benefit analysis considers whether the effort of redaction by the government institution is justified by the benefits of severing and disclosing the remaining information. Even where the severed text is not completely devoid of meaning, severance will be reasonable only if disclosure of the unexcised portions of the record would reasonably fulfill the purposes of the Act.

[175] The Supreme Court cited with approval the following statement by Associate Chief Justice Jerome in *Montana Band of Indians v Canada (Minister of Indian and Northern Affairs)*, 1988 CanLII 9466 (FC), [1989] 1 F.C. 143 (T.D.):

To attempt to comply with section 25 would result in the release of an entirely blacked-out document with, at most, two or three lines showing. Without the context of the rest of the statement, such information would be worthless. The effort such severance would require on the part of the Department is not reasonably proportionate to the quality of access it would provide. [Emphasis added; pp. 160-61.]

[176] Applying this test to the case at bar, the question is whether the “effort” required to further sever the records is “reasonably proportionate to the quality of access it would provide.”

[177] I am not persuaded by the Information Commissioner's assertion that the reference in the French version of section 25 to "problèmes sérieux" sets a different and more rigorous standard than the English version. "Reasonable severance" can involve overcoming problems and expending effort, as is clear from the cost-benefit test set by *Merck Frosst*. It is only where the expenditure of effort is disproportionate to the quality of access that disclosure becomes unreasonable. In my view, the test from *Merck Frosst* is applicable, and no further elaboration is required.

[178] Health Canada submits that it fulfilled its obligations by disclosing the first character of the relevant FSAs, and by considering whether to release more information. It argues that requiring it to perform the type of risk analysis described by the expert for each of the over 1,000 FSAs included in the data here would be an unreasonable burden.

[179] In one respect, I am persuaded that the expert's report provides a blueprint to Health Canada for conducting the type of analysis that is required. I reject Health Canada's argument that it would have to start from scratch in developing its approach to conducting the required risk analysis, because it can start with the code already developed and applied by its expert. On the other hand, I also reject the argument that Health Canada can simply use the expert's code to conduct the analysis on all of the datasets.

[180] The expert's report states that in order to ensure consistency, "specific software will need to be developed to perform the necessary calculations since the analysis needed cannot practically be done manually" (Expert Report, p. 15, R Record p. 979). The expert sets out the

process that would need to be followed, including: identifying the quasi-identifiers (the expert assumes that age, sex, and FSA are included); applying the relevant response knowledge assumption (i.e. does the person know someone in the dataset or not); considering whether any dataset equivalence class represents more than 70% of the population for any particular dataset release (if so, the risk is too high); ensuring that the data is grouped appropriately to manage the risks; and considering whether other available information points to other quasi-identifiers. By any measure, this is not a simple exercise.

[181] In assessing whether the effort is reasonably proportionate to the quality of access, two points should be emphasized: first, the sensitive nature of the information suggests that the lowest-risk option should be adopted; second, the first character of the relevant FSAs has already been disclosed, and so the general location of most of the licenses has been revealed. The question is whether a further narrowing of the lens would bring significant benefits, given the effort that doing so would require.

[182] I find that imposing such a requirement on Health Canada, in the context of the particular facts of this case, goes beyond what section 25 requires.

[183] First, there is no evidence that Health Canada has the “in-house” expertise that is required to carry out this task. The fact that it hired an outside expert for this case is an indication that it does not. I accept Health Canada’s assertion that computer-coding expertise is not sufficient; the type of analysis that is required also involves specialized expertise to conduct an assessment of

the risk factors, including quasi-identifiers, the size and make-up of the equivalence class, and the other available information that is relevant to assessing the privacy risks.

[184] Furthermore, the type of technical considerations set out in the expert's report suggest that the computer program would need to be adapted to consider newly released elements over time. That is to say, the program would have to evolve to keep pace with new risks as they are identified, as further releases are considered, and other information emerges over time – whether from Health Canada releases, other publicly available data, or research on the subject. It is not a static exercise.

[185] Finally, it is not an “all or nothing” proposition, because Health Canada has already released the first character of the relevant FSAs. Some information is already in the public domain, and to the extent that further information could be disclosed, I am persuaded that doing so would entail privacy risks that are too high.

[186] The expert's report does not support the view that some of the high risk FSAs could be further sub-divided without increasing the risk, and it bears repeating that the expert's analysis did not take age and gender into account in the relevant parts of the analysis, but rather focused only on the population of the FSA. The expert has identified age (or age range) and gender as “definitive quasi-identifiers” because of the prior releases, and the expert states that if age and gender are added “the equivalence class sizes would be lower (and hence risk levels would be higher)” (Expert's Report, p. 17).

[187] I have already found that age and gender are relevant quasi-identifiers, so any further analysis would only increase the risks beyond their already unacceptable levels. This also supports the view that further parsing of the datasets is not reasonable.

[188] For similar reasons, I also reject the Privacy Commissioner's submission that Health Canada should apply other de-identification techniques in order to disclose more of the information. I accept that this is an approach that should be considered by a government institution in discharging its obligations under section 25. On this, I am not persuaded by Health Canada's argument that it is forbidden from doing this by section 67.1(1) of the *ATIA*. However, I am also not persuaded that an analysis of other de-identification techniques is significantly less complex than the risk analysis described in the preceding paragraphs. At the end of the day, Health Canada would need to understand how the other techniques successfully lowered the re-identification risks sufficiently to make them a viable alternative, and given the complexities of conducting such an analysis in the circumstances of this case, I find it goes beyond what is required by section 25.

[189] For all of these reasons, I find that Health Canada was not required to undertake further severance in order to meet its disclosure obligations under section 25.

V. Conclusion

[190] For the reasons set out above, this application is dismissed.

[191] It must be acknowledged that there are important interests on both sides of the debate, as the parties and intervener plainly demonstrated. Access to information regarding the administration of the medical marijuana regime raises an important question of public policy, and more transparency around that must be presumed to be a public good. On the other hand, protecting the privacy of individuals with licenses to grow medical marijuana, or who are designated to do so, is also a public good. In addition, the proper analytical approach to assessing privacy risks in relation to disclosure of structured datasets is a novel and important question.

[192] In the end, the jurisprudence, combined with the evidence Health Canada produced, leads me to conclude that the application cannot succeed. The Supreme Court of Canada has made it clear that in a clash between access to information and individuals' privacy rights, privacy must prevail. That is also Parliament's intention, as is evident from the relationship between the *ATIA* and the *Privacy Act*.

[193] On the evidence here, I am persuaded that the risks to privacy that would arise from any further disclosure of the records are simply too great. The evidence demonstrates a serious possibility that disclosing further data about the FSAs and/or city names would risk exposing very sensitive information about individuals, and thus Health Canada's refusal to accept the Information Commissioner's recommendation is justified.

[194] Similarly, I find the evidence compels the conclusion that requiring Health Canada to undertake a risk analysis for each FSA separately would impose a burden on it that is disproportionate to the quality of additional access it would provide.

[195] No party sought costs, and in view of the fact that the parties to the litigation are all public institutions, and the case raised novel and important questions, no costs will be awarded. Each party will be responsible for their own costs.

[196] Finally, I acknowledge and thank counsel for the parties and the intervener for the quality of their written and oral submissions.

[197] A copy of the judgment and reasons will be placed in each of the files.

[198] Postscript: A confidential version of the decision was released to the parties because the record contained some confidential information. The public version takes into account the parties' submissions on any necessary redactions and corrections.

JUDGMENT in T-641-20, T-645-20 & T-637-20

THIS COURT'S JUDGMENT is that:

1. This application is dismissed.
2. Each party will be responsible for their own costs.
3. A copy of the Judgment and Reasons will be placed in each of the Court files.

"William F. Pentney"

Judge

FEDERAL COURT

SOLICITORS OF RECORD

DOCKETS: T-641-20, T-645-20, T-637-20
STYLE OF CAUSE: PATRICK CAIN v MINISTER OF HEALTH AND THE
PRIVACY COMMISSIONER OF CANADA AND
MOLLY HAYES v. MINISTER OF HEALTH AND
THE PRIVACY COMMISSIONER OF CANADA

PLACE OF HEARING: OTTAWA, ONTARIO

DATE OF HEARING: FEBRUARY 7, 2022

JUDGMENT AND REASONS: PENTNEY J.

DATED: JANUARY 13 ~~25~~, 2023

APPEARANCES:

Aditya Ramachandran FOR THE APPLICANTS
PATRICK CAIN AND MOLLY HAYES

Sharon Johnston FOR THE RESPONDENT
MINISTER OF HEALTH

Regan Morris FOR THE INTERVENER
PRIVACY COMMISSIONER OF CANADA

SOLICITORS OF RECORD:

Office of the Information FOR THE APPLICANTS
Commissioner PATRICK CAIN AND MOLLY HAYES
Ottawa, Ontario

Attorney General of Canada FOR THE RESPONDENT
Ottawa, Ontario MINISTER OF HEALTH

Office of the Privacy FOR THE INTERVENER
Commissioner PRIVACY COMMISSIONER OF CANADA
Ottawa, Ontario