

Information and Privacy Commissioner,
Ontario, Canada



Commissaire à l'information et à la protection de la vie privée,
Ontario, Canada

PHIPA DECISION 243

Complaint HI22-00029

University of Toronto

May 21, 2024

Summary: The Information and Privacy Commissioner of Ontario (the IPC) received an anonymous complaint from a group of doctors relating to two research databases created from personal health information, UTOPIAN and POPLAR. The complaint alleged that the personal health information used to populate these databases was obtained from health information custodians without patient consent, and without providing sufficient information to the custodians. The complaint raised concerns about the de-identification of personal health information, and the possibility that such information was being sold or otherwise provided to third parties. The complainants contended that the underlying activity of operating a database of the nature of UTOPIAN and POPLAR was not “research” as contemplated by section 44 of the *Personal Health Information Protection Act* (the *Act* or *PHIPA*), and further alleged that even if this was research, the databases did not otherwise meet the requirements of section 44.

The IPC contacted the University of Toronto (the University), the operator of UTOPIAN. The University stated that POPLAR had been taken over by Queen’s University and was not yet operational, which Queen’s independently confirmed to the IPC. Accordingly, this investigation in respect of the University focuses on UTOPIAN only.

The University provided the IPC with extensive documentation regarding the operation of UTOPIAN. During the course of the investigation, the University stated that it continued to operate its database pursuant to a Protocol Completion Report approved by the University of Toronto Research Ethics Board (REB) after the REB approval for its research plan had expired. Later in this investigation, the University reported that it had paused all UTOPIAN activities and was in the process of applying for REB approval for use of the archived UTOPIAN database for research purposes.

In this decision, I find that the University collected personal health information without authorization under the *Act* during two periods when its REB approval had lapsed. I also find that the University failed to comply with requirements in s. 44 of the *Act* in that it failed to provide health information custodians copies of the research plan and its approval decision, failed to make regular site visits as required under the applicable research plan, and failed to provide custodians with notice of the 2018 collections of personal health information that occurred without an REB approval in place. Finally, I find that the University did not amend its research agreements by merely sending custodians notice of its proposed changes, and to the extent that the University collected, used, and retained personal health information beyond what was permitted by the applicable research agreement, this collection, use, and retention contravened section 44 of the *Act*.

The IPC did not find any evidence to substantiate the complainants' allegations regarding the sale of personal health information, or their de-identification concerns. However, I recommend that in its new application to the REB relating to UTOPIAN data, the University should update its means of notifying patients regarding the UTOPIAN project, conduct a re-identification study to assess the robustness of its de-identification procedures, and exercise greater transparency with contributing custodians. I also recommend that the University ensure that it has research agreements in place with contributing custodians, including any significant amendments hereto, and that it complies with the applicable research agreements.

Statutes Considered: *Personal Health Information Protection Act, 2004, S.O. 2004, c. 3. Sched A.*, sections 2, 3, 4, 44, and 45; *Reg. 329/04*, sections 15 and 18(4).

COMPLAINT:

[1] On July 14, 2022, the IPC received a complaint under the *Personal Health Information Protection Act* (the *Act* or *PHIPA*). This complaint said that it was from a group of doctors who wished to remain anonymous due to whistleblowing concerns.

[2] The complaint was made against the University of Toronto's Practice-Based Research Network (UTOPIAN) and the Primary Care Ontario Practice-based Learning and Research Network (POPLAR). Both UTOPIAN and POPLAR are databases operated by research networks. When referring to UTOPIAN or POPLAR throughout this decision, I am referring to both the relevant research network and database of each. UTOPIAN is operated by the Department of Family & Community Medicine at the University of Toronto (the University).

[3] The anonymous complainants described themselves as physicians with complaints about UTOPIAN arising from their own particular experiences and their concerns for patients within Ontario.

[4] The complaint stated that the UTOPIAN database (also called Data Safe Haven by the University) contained patient medical data obtained from physicians in the community and hospitals, and that the University did not seek patients' consent prior to including their personal health information in this database. Instead, physicians entered into

agreements directly with the University to supply this data to populate the UTOPIAN database. The complaint questioned whether the University gave participating physicians enough time or information for the physicians to provide informed agreement for the use of this personal health information. The complaint noted that patients were largely unaware of the database project, so were not afforded a real opportunity to withdraw from it.

[5] The complaint stated that in 2020, some years into the project, the University significantly increased the extent of the information it uploaded from physicians' electronic medical records (EMR) systems. This uploaded information was expanded to include the entirety of patients' charts and identifying patient information. The complaint stated that the University sent out a letter alerting at least some physicians to this change but did not provide UTOPIAN's research plan or documentation showing that the University of Toronto's Research Ethics Board (the REB) had approved this increase in the amount and nature of data collected.

[6] The complaint raised concerns about UTOPIAN keeping linkage files, which would allow personal health information to be re-identified in the future, and raised the possibility that identifying information may be leaked. Finally, it raised the possibility that the University may be selling or otherwise providing the personal health information it obtains from physicians to third parties.

[7] At the core of the anonymous complaint is the contention that the operation of the UTOPIAN database is not research as allowed under *PHIPA*, and that the University therefore needed patients' consent before their personal health information could be uploaded to this database. The complaint further states that even if the operation of UTOPIAN was research, the University did not meet the requirements set out for researchers found in section 44 of the *Act*.

[8] Because the complaint was filed anonymously and this office had no way of contacting the complainants, the IPC opened this Commissioner-initiated complaint file to investigate the allegations.

BACKGROUND:

The UTOPIAN project

[9] UTOPIAN is one of several Practice Based Learning and Research Networks (PBLRNs) in Canada. The University describes the role of such organizations as follows:

Through the collection of information from primary care electronic medical records (EMRs), PBLRNs provide a platform for epidemiological surveillance, clinical research in practices and data-informed quality improvement relevant to the context of practices and communities. EMRs contain data on acute symptoms as patients reach out to their providers, information on

underlying chronic conditions, mental health, vital signs, social determinants of health and risk factors for disease. They are large and rich sources of population-level data in our communities.

[10] UTOPIAN was established in 2013, when the REB approved the project's Ethics Review Protocol Submission Form. The University described UTOPIAN as a research database composed of de-identified records of approximately 600,000 patients extracted from EMRs of nearly 600 contributing primary care practices, in and around Toronto.

[11] The University has described the information UTOPIAN collected from these contributing practices as personal health information. "Personal health information" is defined under section 4(1) of *PHIPA* to encompass identifying information about an individual relating to their physical or mental health, information relating to the provision of health care to an individual, and the individual's health card number, among other things. (See Appendix 1 for the full text of section 4(1) of *PHIPA*)

[12] The University states that UTOPIAN's authority to collect and use personal health information is found in section 44 of *PHIPA*, which permits health information custodians (such as the primary care providers) to disclose personal health information to researchers if certain conditions are met. (See Appendix 1 for the full definition of "health information custodians" under section 3(1) of *PHIPA*.)

[13] The University states that in addition to providing de-identified data from the UTOPIAN database to outside researchers for specific projects, the UTOPIAN project is itself a research project approved by the REB.

[14] The University states that patient consent is not needed to use personal health information for research purposes under *PHIPA*. Rather, section 44 requires that the researcher have in place a REB-approved research plan and an agreement with the contributing health information custodians for use of their patients' personal health information. The University states that under the terms of its research plan¹, it is the custodians contributing information to UTOPIAN who must provide notice to patients that their personal health information is being used in this database and inform these patients that they can opt out of this use. If a patient opts out, the University removes their data from the research database, though it is not able to remove it from datasets that have previously been provided to outside researchers or from research that has already been conducted using their personal health information.

How UTOPIAN worked

[15] The UTOPIAN database operated by extracting personal health information from

¹ As set out in paragraph 36, there were a series of research plans submitted to the REB as either renewals or amendments. Unless the context indicates otherwise, when I use "research plan" I am referring to the last Human Participant Ethics Protocol Submission, dated September 9, 2022, which expired on November 13, 2022.

custodians' EMRs, either via the EMR provider, a third-party agent, or the University itself (acting as the custodian's agent).

[16] The University states that this raw data containing patient identifiers was partitioned off from the database and stored separately in a secure area of the server, with only UTOPIAN data managers having access to the data containing these identifiers.

[17] UTOPIAN then replaced identifying information with proxy variables², and created linkage files to track these replacements. The University states that these linkage files allowed for later re-identification of the data, if required. They also served to recognize if data from subsequent uploads was associated with a patient already within the database. The University notes that the ability to re-identify individuals is necessary for longitudinal studies, as well as to remove the personal health information of individuals who ask to be removed from the database.

[18] The University notes that once the identifying information was replaced by proxy values and safety checks were performed, the de-linked data was placed in the research database. Researchers do not have access to the database itself. Instead, they are given access to research datasets that UTOPIAN staff provide within a separate zone of the database, called SAVE (Secure Analytic Virtual Environment).

[19] Researchers using UTOPIAN data are required to enter into a "One-Way Confidentiality Agreement" (the Researcher NDA). That agreement states that researchers agree to maintain in confidence any information that is "non-public, confidential or proprietary in nature," which it accessed via the UTOPIAN data platform. Under the Research NDA, researchers also agreed to only use that confidential information for the purpose of the specified work and to comply with privacy laws including *PHIPA*. It also states that the researcher will not attempt to re-identify any individuals associated with any confidential information, and that researchers are not permitted to download any data.

[20] Initially, UTOPIAN's research plan specified that it would not extract any direct patient identifiers but would collect "demographic information, as well as conditions, medications, encounters, and other data relevant to chronic and acute conditions."

[21] In 2020, the University expanded the scope of the personal health information UTOPIAN collected to essentially the entirety of patients' medical charts. At that time, the University informed custodians that it would be collecting patient identifiers such as patient names, addresses, phone numbers, emails and health card numbers. Also, UTOPIAN began collecting personal health information from within free text fields, as well as that captured within picture-like formatted documents (generally PDF, JPG, or TIFF). This could include personal health information found in progress notes, diagnostic tests, consult notes, hospital discharge summaries and referral letters.

² This process is set out in more detail in paragraphs 111-116 of this Decision.

[22] To illustrate the types of information collected after this expansion of scope, the University provided the example of a consultation letter sent via fax to a family physician. A consultation letter may be received as a PDF and added to the patient's electronic file. This PDF could include the patient's name and contact information, as well as the patient's presentation at the consultation, tests, diagnosis and recommendations for follow up.

How Custodians Signed Up

[23] The University states that it recruited health information custodians to provide personal health information from their EMRs for the database. In the process of doing so, it provided custodians with "UTOPIAN Data Safe Haven: Letter of Information and Physician Consent" (the Provider Agreement).³ The Provider Agreement is a multi-page letter describing the UTOPIAN project, followed by a custodian consent form. This form asks custodians to check "yes" or "no" to questions relating to their understanding of the project, concluding with a question on whether they agree to participate. If the custodian agreed to participate, they would then sign the document to confirm their understanding of, and agreement to participate in, the project.

[24] The Provider Agreement set out the purpose of UTOPIAN as the creation of "a de-identified, standardized, and researchable database ... using health data from the EMRs of participating primary care providers." It lists the project's objectives but does not explicitly describe UTOPIAN itself as a research project, other than parenthetically mentioning that UTOPIAN has its own research protocol. It does so in the "Potential uses of data in Clinical Research," where the Provider Agreement states that additional REB approval would be needed to recruit physicians and patients for additional research studies, noting that "this will be in addition to the UTOPIAN Data Safe Haven REB approval (this protocol)." The "Questions" section does refer to participation in a "study" and directs those with "questions about your rights as a research participant or how the study is being conducted" to contact the REB.

[25] Within the Provider Agreement, the University describes the process for patient data collection and the partition of identifying information, as well as the de-identification process. It states that the data will be retained for "the duration of the database, which is intended to be ongoing", as long as the patients have not opted out of its use or the custodians have not withdrawn agreement for its use. The Provider Agreement states that custodians will inform patients of the use of their personal health information for the UTOPIAN project, and their ability to opt out of it, via posters displayed in their office waiting areas. It also notes that it is important for custodians to alert patients to these posters.

[26] The Provider Agreement does not cite any section of *PHIPA* as authority for why custodians may disclose personal health information to the University and does not

³ As noted in paragraph 142, UTOPIAN revised the Provider Agreement during the course of that project, resulting in eight different versions of the Provider Agreement. Unless the context indicates otherwise, when I refer to the Provider Agreement, I am referring to the most recent version.

include a copy of any UTOPIAN research plan or its REB approval decision.

UTOPIAN's Research Ethics Board (REB) Approvals

[27] *PHIPA* provides for the disclosure and use of personal health information for research purposes without the patient's consent, if the criteria set out in s. 44 are met. Central to these criteria is that a researcher must submit a research plan to a REB and obtain REB approval for that plan.

[28] Under *PHIPA*, a REB is "a board of persons that is established for the purpose of approving research plans under section 44 and that meets the prescribed requirements." The requirements referred to in the above quote are set out in section 15 of Regulation 329/04 to *PHIPA*, including the required composition of the board:

The board must have at least five members, including,

- I. at least one member with no affiliation with the person or persons that established the research ethics board,
- II. at least one member knowledgeable in research ethics, either as a result of formal training in research ethics, or practical or academic experience in research ethics,
- III. at least two members with expertise in the methods or in the areas of the research being considered, and
- IV. at least one member knowledgeable in considering privacy issues.

[29] The University of Toronto established its own REB pursuant to *PHIPA*.

[30] The University provided the IPC with all its applications to the REB throughout the life of the UTOPIAN project. As part of these, the University completed the REB's required forms and provided supporting documents containing information on the purpose, methodology, proposed participants, experience of the research team, potential risks and benefits, means of participant withdrawal, privacy and security protections, data security protections, and risk level of the project.

[31] The University first sought approval from the REB in 2013 for the UTOPIAN project. That application grew out of the existing Canadian Primary Care Sentinel Surveillance Network (CPCSSN), a Canada-wide electronic medical record surveillance system, of which UTOPIAN represented one project.

[32] The initial submission seeking REB approval listed three purposes of the UTOPIAN project:

- To create a database with anonymized patient data from EMRs of primary health care providers
- To provide accessible data options for research and public health surveillance
- To devise algorithms or other processes to enable automated EMR data collection, data de-identification, and other data processes.

[33] In its submission, the University stated it would extract, clean, code, and upload health data that has been “fully anonymized”. It specified that extraction would take place every three months and described the health data extracted as conditions, medications, encounters, and other data relevant to chronic and acute conditions, as well as demographic data. The application stated that no direct identifying information would be included within the database. The UTOPIAN database would be stored on the CPCSSN server, and the UTOPIAN data added to the CPCSSN central repository.

[34] The University also stated in its application that patient consent was not required for this secondary use of health data, but that patients would be notified of the use via brochures and posters⁴ at the custodian’s office. These would address the patients’ option for withdrawal from the project. Patients who withdrew would no longer have their data extracted and the University would make every effort to remove this data from the CPCSSN central repository.

[35] Finally, the University’s application stated that if the project were to end with no defined and approved successor organization, the data would be destroyed.

[36] Over the next several years, the University applied to the REB for approval of renewals of, and amendments to, the UTOPIAN research plan. During this time, the UTOPIAN research plan underwent several amendments including:

- The addition of a study linking UTOPIAN data to data held by the Institute for Clinical Evaluation Studies (ICES);
- Following the success of the above-noted ICES study, a request that the linkage be ongoing;
- A name change to UTOPIAN Data Safe Haven, due to the project no longer being associated with CPCSSN; and
- Forwarding data to Diabetes Action Canada.

[37] The University proposed a much more significant amendment to the UTOPIAN research plan in 2020. At that time, the University asked for REB approval to extract all information from free text fields from the EMRs, including progress notes, diagnostic

⁴ At that point, these were CPCSSN posters and brochures, rather than UTOPIAN-specific literature.

tests, consult notes, hospital discharge summaries, and referral files. The University described this change, and the reasons for it, as follows:

Essentially, this would include all clinically relevant information contained in the EMR. Some of these elements are stored in pdf (picture like format) and thus cannot be de-identified. We would like to collect these items and store them securely to allow data to be extracted from them for chart abstraction studies. These identifiable data would be stored separately, and would not be available to the regular UTOPIAN staff; they would be available to chart abstractors who sign confidentiality agreements prior to data access. The necessity of having full chart access is to be able to have the full clinic picture for validation studies or other studies that require more detailed clinical information. *Having this unstructured free text data also makes the data more amenable to artificial intelligence/text-mining methodology.* [Emphasis added.]

[38] In addition to the clinical information, the University also requested that they be permitted to collect patient names, addresses, phone numbers, emails, and health cards, stating that this would “improve processes and procedures employed to de-identify both the structured and unstructured text data that we collect.” The University stated that this additional identifying information would be partitioned immediately upon receipt, stored separately from the rest of the data, and only accessed by personnel for the purpose of improving de-identification procedures.

[39] Finally, in that same application, the University also asked that it be permitted to link to administrative and other data from any prescribed entity, rather than limiting that linkage to only ICES.

Relationship between UTOPIAN and POPLAR

[40] In early 2021, the University requested a name change from UTOPIAN to Primary Care Ontario Practice-Based Learning and Research Network (POPLAR) Data Platform, with the shortened title of POPLAR Data Platform⁵. This was done in anticipation of POPLAR becoming operational, building on UTOPIAN’s database and structures as well as those of other similar research networks. This name change from UTOPIAN to POPLAR was approved by the REB on April 1, 2021, and the University intended to apply for separate REB approval of POPLAR.

[41] POPLAR is described by the University as an evolution of UTOPIAN and similar existing networks elsewhere in Ontario:

⁵ As I will later address, UTOPIAN never merged into POPLAR while at the University of Toronto. Given this, I will continue to refer to this project as UTOPIAN and will refer to the larger project that was eventually approved by the Queen’s REB as POPLAR.

[POPLAR] is a central hub and platform that will centralize the data from practice-based learning and research networks across Ontario comprised of six University Departments/Sections of Family Medicine and the Alliance for Healthier Communities (Ontario's Community Health Centres, Nurse Practitioner Led Clinics and Aboriginal Health Access Centres).

[42] The University's plan was that once POPLAR became operational, UTOPIAN would essentially be absorbed into the larger POPLAR network and would form the backbone of this new network. At the time of the requested name change, POPLAR was still in its planning stages and the POPLAR database did not yet exist.

[43] The University initially applied to the REB for approval of POPLAR on June 9, 2022. However, the University subsequently informed me that POPLAR withdrew its application to the University of Toronto REB and instead submitted an application to the Queen's Health Sciences and Affiliated Teaching Hospitals Research Ethics Board (Queen's REB) in September 2023. The Queen's REB granted initial approval to the POPLAR application on November 15, 2023, though as of the time of this writing, the POPLAR database is not yet operational.

[44] Meanwhile, the UTOPIAN REB approval came up for renewal. The University applied for two consecutive UTOPIAN renewals, each for two-month periods.

Events during the course of the IPC Investigation

[45] On October 18, 2022, following receipt of the anonymous complaint, I attempted to contact the individual publicly listed as the Director of UTOPIAN. Shortly thereafter, the University informed me that I could direct all correspondence in this matter to the University itself. I therefore sent the University the Notice of Investigation in this matter on November 1, 2022, which included some preliminary questions.

[46] Based on the University's November 25, 2022 response to the questions in the Notice of Investigation, I determined that it was necessary to commence a review of the operation of UTOPIAN under section 58(1) of *PHIPA*. I sent the University a Notice of Review on January 17, 2023, with detailed questions regarding UTOPIAN's operation.

[47] The University's counsel then sent me a letter on January 25, 2023, informing me that the University had recently discovered that UTOPIAN's REB approval had expired as of November 13, 2022, and that "some automated data extraction continued" on January 5, 2023, some two months after the REB approval had lapsed.

[48] More specifically, the University discovered that UTOPIAN had collected personal health information from four primary care groups comprised of 29 family physicians on January 5, 2023. The University stated that the vendor extracted this data because the upload had been previously scheduled, and the University did not pause the automated data extraction after the REB approval lapsed.

[49] When it found out about the upload, the University immediately checked for any other upcoming automatic uploads and cancelled these. It stated that it did not process or use the personal health information that it collected on January 5, 2023 and subsequently conducted an audit to confirm there had been no access to this data.

[50] On January 20, 2023, the University notified the REB of the January 5, 2023 collections by submitting a Protocol Deviation Report. This report stated that the data was not accessed, processed, or used in any way, and that the server had been shut down to prevent further uploads. The University proposed to restart the server to destroy the data, and then immediately shut it down again, a process which the REB approved. The data was subsequently destroyed, with the University providing a Certificate of Destruction to confirm this action.

[51] On January 25, 2023, the University notified the affected custodians of the "inadvertent upload of personal health information" that was "inconsistent with PHIPA".

[52] Shortly after, on February 8, 2023, the University submitted a Protocol Completion Report to the REB. The reason provided was that the study in its current form was completed and would no longer continue to extract new data from EMRs. The University planned to join POPLAR instead, with the UTOPIAN data to be later transferred to POPLAR.

[53] The University stated that if the REB approved the Protocol Completion Report, UTOPIAN would then be a database only, under which researchers who have REB-approved protocols would be provided with research data. The REB approved this Protocol Completion Report, and UTOPIAN proceeded to operate in this manner.

[54] After receiving this information, I sent an Amended Notice of Review on February 23, 2023, which included additional questions. The University sought a time extension, which I granted to April 14, 2023. At that time, the University provided a thorough response to my inquiries, providing extensive documentation related to the REB approval process. This included all UTOPIAN applications to the REB, the Provider Agreement, the Researcher NDA, the Patient Information Poster, de-identification procedures, and the Protocol Deviation Report. The University also provided me with the Protocol Completion Report stating that UTOPIAN had stopped extracting new data from EMRs and would continue to operate as a database only.

[55] After reviewing the University's response and the associated documentation, I sent the University follow-up questions on November 17, 2023. In that letter, I also informed the University that pursuant to section 11 of *PHIPA Practice Direction #3: Publicly Released Decisions under the Personal Health Information Protection Act, 2004*,⁶ and given the public interest in this matter, the University should expect to be named in any decision

⁶ <https://www.ipc.on.ca/resource/hipa-practice-direction-3-publicly-released-decisions-under-the-personal-health-information-protection-act-2004/>

that I release.

[56] Upon seeking and receiving an extension, the University provided a response to my questions on December 20, 2023, confirming that it has indeed ceased data collection and that it continues to operate as a database only. The University stated that previously approved research projects continued to use data from UTOPIAN and that it had the authority to provide new projects with research data, subject to REB approvals of those specific projects.

[57] In April 2024, I sent the University some additional follow up questions that the University responded to that same month. In their response, the University stated that “[as] of February 2024, all activities related to the UTOPIAN database have been paused.” The University also stated that in consultation with the REB it had decided to make a new application to the REB for the continued use of the archived UTOPIAN database for research purposes.

Relevant legislation

[58] The University relies on the provisions in section 44 of *PHIPA* for its authority to collect personal health information from custodians, and then use that information for the purposes of research. *PHIPA* defines the terms “research” and “researcher” at section 2 as follows:

“research” means a systematic investigation designed to develop or establish principles, facts or generalizable knowledge, or any combination of them, and includes the development, testing and evaluation of research;

“researcher” means a person who conducts research;

[59] Section 44 permits health information custodians to disclose personal health information for research purposes if certain conditions are met. The portions of section 44 relevant to the current investigation are:

Disclosure for research

44 (1) A health information custodian may disclose personal health information about an individual to a researcher if the researcher,

(a) submits to the custodian,

(i) an application in writing,

(ii) a research plan that meets the requirements of subsection (2),
and

(iii) a copy of the decision of a research ethics board that approves the research plan; and

(b) enters into the agreement required by subsection (5).

[...]

Research plan

(2) A research plan must be in writing and must set out,

(a) the affiliation of each person involved in the research;

(b) the nature and objectives of the research and the public or scientific benefit of the research that the researcher anticipates; and

(c) all other prescribed matters related to the research.

[...]

Decision of board

(4) After reviewing a research plan that a researcher has submitted to it, the research ethics board shall provide to the researcher a decision in writing, with reasons, setting out whether the board approves the plan, and whether the approval is subject to any conditions, which must be specified in the decision.

Agreement respecting disclosure

(5) Before a health information custodian discloses personal health information to a researcher under subsection (1), the researcher shall enter into an agreement with the custodian in which the researcher agrees to comply with the conditions and restrictions, if any, that the custodian imposes relating to the use, security, disclosure, return or disposal of the information.

Compliance by researcher

(6) A researcher who receives personal health information about an individual from a health information custodian under subsection (1) shall,

(a) comply with the conditions, if any, specified by the research ethics board in respect of the research plan;

(b) use the information only for the purposes set out in the research plan as approved by the research ethics board;

(c) not publish the information in a form that could reasonably enable a person to ascertain the identity of the individual;

(d) despite subsection 49 (1), not disclose the information except as required by law and subject to the exceptions and additional requirements, if any, that are prescribed;

(e) not make contact or attempt to make contact with the individual, directly or indirectly, unless the custodian first obtains the individual's consent to being contacted;

(f) notify the custodian immediately in writing if the researcher becomes aware of any breach of this subsection or the agreement described in subsection (5); and

(g) comply with the agreement described in subsection (5).

PRELIMINARY MATTERS:

[60] While the complaint raised concerns with both POPLAR and UTOPIAN, the focus in this decision is on UTOPIAN. POPLAR was not active at the time of the complaint and remains inactive at the time of writing this report.

[61] There is no dispute, and I find, that the information uploaded to the UTOPIAN database from custodians was personal health information as defined under section 4(1) of the *Act*.

[62] There is no dispute, and I find, that the personal health information uploaded to the UTOPIAN database was collected from health information custodians.

[63] The concerns set out in the complaint were about the operation of the UTOPIAN database which is operated by the University. As such, the University is the focus of, and the respondent to, this investigation, rather than the individual health information custodians who contributed personal health information to UTOPIAN. In this decision, I provide no comment on whether custodians complied with *PHIPA* in the context of their disclosures of personal health information to UTOPIAN, recognizing that they may have individualized facts and circumstances beyond the scope of this investigation.

[64] The University has stated that it does not rely on the consent of patients to collect personal health information pursuant to section 44(1), as that section does not require the consent of patients for health information custodians to disclose their personal health information to researchers. The University takes the position that it is a "researcher" conducting "research" under ss. 2 and 44 of *PHIPA*. For the purposes of my examination of compliance with the individual provisions of section 44, I will assume that to be true, but will address that claim later in the decision.

ISSUES:

1. Assuming UTOPIAN is research under the *Act*, did the University comply with the requirements of section 44 of *PHIPA*?
 - a. Were custodians provided with a research plan and REB approval pursuant to section 44(1)(a) of the *Act*?
 - b. Did custodians enter into an agreement pursuant to section 44(5) of the *Act*?
 - c. Did the University comply with the conditions, if any, specified by the research ethics board in respect of the research plan pursuant to section 44(6)(a) of the *Act*?
 - d. Did the University use the personal health information provided only for the purposes set out in the research plan as approved by the research ethics board pursuant to section 44(6)(b) of the *Act*?
 - e. Did the University provide notice to custodians of a breach of section 44(6) of the *Act*, or a breach of the research agreement, pursuant to section 44(6)(f) of the *Act*?
 - f. Did the University comply with the research agreement as required under section 44(6)(g) of the *Act*?
2. Is the creation of the UTOPIAN database “research” pursuant to sections 2 and 44 of the *Act*?
3. Should this matter proceed to adjudication at the IPC, where a potential order may be issued?

DISCUSSION:

1. Assuming UTOPIAN is research under the *Act*, did the University comply with the requirements of section 44 of *PHIPA*?

a. Were custodians provided with a research plan and REB approval pursuant to section 44(1)(a) of the Act?

[65] Section 44(1)(a) clearly states that disclosure to a researcher is conditional on that researcher submitting both a research plan (containing specified elements) and a copy of the REB decision approving that plan to the custodian.

[66] The University did not provide a copy of the research plan to custodians who

contributed patient data to UTOPIAN when they signed up to UTOPIAN, and likewise did not provide the updated research plans following renewals and amendments to that plan. Instead, it gave custodians the Provider Agreement, which consists of a letter outlining the UTOPIAN project and a consent form for the custodian to sign prior to contributing personal health information for the project. The University set out its rationale for doing so as follows:

The University designed the "Letter" to be readable to its audience. It serves as a form of application, setting out the approved research plan and setting out e-mail and phone contact information for the research ethics board with an invitation to ask questions. The University made copies of research ethics board approvals available to health information custodians upon request.

[67] Although the University takes the position that UTOPIAN is authorized to collect personal health information under the terms of section 44 of *PHIPA*, the Provider Agreement is largely silent on the fact that UTOPIAN is itself a project with its own research plan and subject to REB approval, other than the parenthetical references mentioned in paragraph 24 above.

[68] The Provider Agreement is also silent on the duration of the UTOPIAN database, and that its upload of data is contingent on ongoing REB approval, which would have put custodians on notice of the inherent possibility that subsequent approvals may not be granted, or may lapse, as occurred in this case. In choosing to provide a less technical, more "readable" document to custodians, rather than the research plan and REB decision approving that plan specified under *PHIPA*, the University failed to inform custodians of pertinent information – namely, that UTOPIAN was a project whose authority for uploading EMR data could end. While I understand, and commend, the University's effort to provide information in more digestible format, it could have easily given the custodians the Provider Agreement as well as the current research plan along with the REB decision approving that plan, letting them know that the research plan contained more detailed information if they wished to refer to it.

[69] The choice the University made limited the information that custodians were given about UTOPIAN generally, as well as the significance of subsequent amendments, particularly the major change to full chart extraction. It also resulted in the University failing to comply with section 44(1)(a) of *PHIPA*, which requires that researchers provide custodians with both the research plan and a REB decision approving the research plan.

[70] These documents are provided for a purpose. Among other things, they confirm that the REB has granted approval to a project and specify its approval period. Providing them to custodians, at minimum, informs them that the project remains ongoing, and when the approval could end.

[71] While the University states it would provide copies of the REB approval if custodians requested them, the Provider Agreement does not mention this. The

University confirmed that it did not “expressly inform” custodians that they could access these. The University notes that no custodian has requested these documents, and attributes this to the digestible information it provided via the Provider Agreement. The University’s logic on this point does not withstand scrutiny, as it appears custodians did not even know they could ask for them.

[72] Moreover, implicit in the requirement to provide custodians with a copy of the REB approval decision in accordance with s. 44 of *PHIPA* is that the REB approval must continue to authorize a disclosure at the time the disclosure is made (e.g. the REB approval must not have expired).⁷

[73] In this case, the University collected personal health information from custodians after the REB approval for UTOPIAN expired. The University had pre-scheduled uploads from EMRs and failed to cancel these after the REB approval expired. The uploads continued to occur as scheduled, resulting in disclosures of personal health information from multiple custodians at a time when there was no REB-approved research plan in place. Those uploads were clearly not authorized under *PHIPA* at that time. The University acknowledged this in its January 25, 2023 letter to the affected custodians, in which it stated that “this inadvertent upload of personal health information is considered to be inconsistent with *PHIPA*.”

[74] The University did not incorporate the uploaded information into the UTOPIAN database, and subsequently destroyed it. Be that as it may, the fact that the University continued to collect personal health information without authorization is a significant breach of the University’s obligations under section 44(1)(a) of the *Act*. At the time of the January 2023 collections, no REB-approved research plan existed, as it had lapsed months earlier. Without providing custodians with the research plan and without there being a valid REB decision approving that plan, the collection of personal health information was not in compliance with section 44(1)(a) of the *Act*.

[75] The University was clear that it was the party responsible for the collections, not the custodian. The University specifies in both its research plan and the Provider Agreement that the UTOPIAN Data Analytics Manager is responsible for overseeing data extraction. Moreover, the University states that it was aware of the expiry date of the REB approval. Despite this, the University did not cease automated extractions upon the expiry date.

[76] After learning of the unauthorized collection, I asked the University if it had notified the Canadian Institutes of Health Research (CIHR), Canada’s federal funding agency for health research, of this lapse in REB approval under the terms of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2 (2022)*. The University informed me that it had not done so. While the University stated that it did not receive CIHR funding for UTOPIAN, the University of Toronto is a CIHR-funded institution,

⁷ For the sake of simplicity, I refer to this as the REB approval remaining “valid” in this decision.

and a number of CIHR-funded projects used UTOPIAN datasets in their research.⁸

[77] In addition, during the investigation of this matter it came to light that this was not the first time the University allowed the REB approval to lapse while continuing to collect personal health information into UTOPIAN. On a previous occasion, the REB approval for UTOPIAN expired after August 6, 2018. Yet, the University did not apply for REB renewal before that approval expired, with the result that the UTOPIAN project operated absent REB approval from August 7, 2018 to October 4, 2018. During this lapse, the University collected personal health information from 33 custodians via scheduled quarterly extractions.

[78] The University explained that this earlier lapse was due to an error within the electronic management system it uses to manage REB applications, which failed to send the expected notifications. The University states that it submitted a renewal application to the REB as soon as it discovered this lapse, and as of October 5, 2018, REB approval was again in place. The October 5, 2018 renewal application did not mention that collections took place when no REB approval was in place. The University did not provide custodians with notice of this unauthorized collection of personal health information, and the personal health information was not removed from the UTOPIAN database at that time.

[79] The University has acknowledged, and I agree, that it did not have authority under section 44 of the *Act* to collect personal health information during the two months in 2018 when no REB approval was in place.

[80] While both the 2018 and 2023 lapses may seem like mere administrative oversights, proceeding to collect personal health information without lawful authorization is deeply concerning from both a legal and ethical perspective. It is also concerning that the University's internal checks did not prevent these lapses.

[81] In summary, by failing to provide custodians with copies of the research plan, and copies of the REB approval, I find that the University failed to meet the requirements set out in section 44(1)(a) of the *Act*. I also find that the University collected personal health information without authority under section 44(1)(a) of the *Act* for the collections that occurred on January 5, 2023 and between August 7, 2018 to October 4, 2018.

b. Did custodians enter into an agreement pursuant to sections 44(1)(b) and 44(5) of the Act?

[82] As a further condition for custodians to disclose personal health information to UTOPIAN for research purposes, sections 44(1)(b) and 44(5) of *PHIPA* require the

⁸ See for example, a study relating to childhood immunizations, available at https://webapps.cihr-irsc.gc.ca/decisions/p/project_details.html?applId=460267&lang=en and a study relating to the use of an app in cardiovascular disease prevention, available at https://webapps.cihr-irsc.gc.ca/decisions/p/project_details.html?applId=458261&lang=en

researcher to “enter into an agreement with the custodian in which the researcher agrees to comply with the conditions and restrictions, if any, that the custodian imposes relating to the use, security, disclosure, return or disposal of the information.”

[83] The University states that the Provider Agreement was the research agreement for the purposes of section 44(5) of the *Act*. As noted, this document is a letter outlining the UTOPIAN project, followed by a one page “Consent Form.” This form has the custodian confirm that they understand various aspects of the project, including that UTOPIAN will have access to the records of their patients, and that the custodian can withdraw at any time. In addition to that confirmation, the Consent Form also states that copies of the patient data extracted from their practices will be forwarded to listed agencies⁹ and then gives the custodian the opportunity to check off any organization that they do not agree to forward a copy on to. Finally, the form asks the custodian to check yes or no to the statement “I agree to take part in this study” and sign the form. Through the lifetime of the UTOPIAN project, there were eight versions of the Provider Agreement, all of which followed the same format of an informational letter followed by a consent form.

[84] The Provider Agreement does not refer to itself as a research agreement, though it does ask that custodians read the letter before completing the included form. The University states that the Provider Agreement communicated that it is a research agreement, noting that it refers to UTOPIAN as a study and identified participating custodians as “research subject(s).” The University also noted that it includes a reference to “the UTOPIAN Data Safe Haven REB approval (this protocol).”

[85] The letter portion of the Provider Agreement sets out the purpose of the project, how UTOPIAN obtains and protects personal health information (including by de-identifying it), what the data may be used for, and the procedures that it will take if a breach occurs. The earliest versions of the Provider Agreement, dating from the time it was still referred to as a CPCSSN project, state that the project includes developing a database; later versions do not. The consent form portion of the Provider Agreement takes steps to verify that the custodian understands what the personal health information will be used for and gives the custodian an option as to whether the information from their practice would be forwarded on to other entities. The custodian can then sign the form to indicate their agreement.

[86] The University states that not all custodians signed on to the Provider Agreement “as is,” noting that individualized agreements were also possible. Some health information custodians requested special restrictions (one such example was not disclosing health card numbers to ICES), and the University stated that eight contributing custodians entered into data transfer agreements in addition to signing the Provider Agreement. The University states that it tracked agreements for compliance with conditions and

⁹ The agencies the Provider Agreement lists are:

1. The Institute for Clinical Evaluative Sciences (ICES) or other prescribed entity
2. The Canadian Primary Care Sentinel Surveillance Network (CPCSSN)
3. Diabetes Action Canada (DAC)

restrictions.

[87] While the more recent versions of the Provider Agreement are certainly not as clear as they could be in describing UTOPIAN itself as a research project, overall I am satisfied that custodians would be broadly aware in signing this agreement that they are agreeing to a disclosure to the University for research purposes in accordance with its terms. The Provider Agreement does inform the custodian what personal health information is being collected, what will happen to it, and what purposes it will ultimately be used for. It is a document that is centred around research and asks physicians to agree to disclose patient personal health information for research purposes. Based on this, I find that the Provider Agreement, when signed by a custodian, is a research agreement as contemplated by sections 44(1)(b) and 44(5) of the *Act*.¹⁰

[88] Having determined that the signed Provider Agreements are research agreements as contemplated under the *Act*, this leaves the question of whether the University successfully amended the Provider Agreement when it provided emailed correspondence informing custodians of significant changes to its research plan.

[89] None of the versions of the Provider Agreement addresses how it may be amended. When asked whether the Provider Agreement was amended during the course of the UTOPIAN project, the University provided the following responses:

The Letter of Information and Physician Consent remained the agreement. It changed over time following research ethics board amendment submissions and approvals. Newly recruited custodians signed the current version of the Letter of Information and Physician Consent.

UTOPIAN notified contributing custodians of every major change. These changes included, for example, sharing with ICES, sharing to the DAC Repository and the move to full chart extraction. Notification was by way of e-mail (as approved by the research ethics board), and UTOPIAN also updated the copy of the Letter of Information and Physician Consent posted on its website. Contributing custodians were not required to sign a new letter.

[90] The University also stated that it sent emails to custodians, describing the changes. The University noted that it “believed contributing custodians needed to learn about the changes, and it elected to send a concise and clear communication to do so rather than the entire amended Letter of Information and Physician Consent.”

[91] One of these changes was the move to full chart extraction, as communicated to custodians in a June 12, 2020 email from the UTOPIAN Director (the Amendment Email).

¹⁰ This finding only applies to the version of the Provider Agreement and accompanying consent form originally signed by the custodian. I recognize that, in a small number of cases, custodians entered into individualized agreements with the University. In this decision, I make no comment on those agreements.

This email described the expanded collection as follows:

[W]e will be adding free text fields such as progress notes, diagnostic tests, consult notes, hospital discharge summaries and referral letters. Essentially, this would include all clinically relevant information contained in the EMR. We will collect these items, de-identify them and store them securely. Some of these data elements are captured in the EMR in a picture like format ie. pdf or tiff.

[92] The Amendment Email also states that UTOPIAN would begin collecting patient identifiers that would be partitioned immediately and stored separately. It states that the REB reviewed and approved this amendment and asked custodians to let the University know if they have any questions or concerns. It concludes by stating "if you do not wish to participate in the amended process, please let us know and we will not extract any additional information." The Amendment Email did not include a consent form for custodians to agree to, instead placing the onus on custodians to opt-out. This is particularly concerning given that the University's stated purpose for moving to unstructured free text data was in part to "make the data more amenable to artificial intelligence/text-mining methodology."

[93] To alert custodians about the expanded chart access, the University sent the Amendment Email, and followed up with an additional email two weeks later, using the "read receipt" function for both. The UTOPIAN Research Coordinator contacted any practices from whom they received a bounce back email.

[94] As noted, the University did not provide a later version of the Provider Agreement to be executed by custodians who had already executed an earlier version of it. Instead, the University took the position that when the REB approved an increase to the scope of data collection, it "acted reasonably to advise contributing custodians, invite questions and opt outs, and amend the agreement." The University states that custodians' failure to opt out after receiving the email correspondence constitutes acceptance of revisions to the Provider Agreement, stating as follows:

There is no legal requirement to amend a section 44 agreement via formal "execution" and no legal requirement for a "positive affirmation" nor can such requirements be read into PHIPA. The purpose of the agreement requirement in section 44 is to facilitate research while supporting contributing health information custodians' control over their contribution of personal health information. This purpose can be met without creating the impediment to ongoing research associated with seeking express amendments. Effective communication is the requirement, and the evidence ... establishes that the University effectively conveyed the changes and obtained the agreement required by PHIPA.

[95] Section 45(5) states that, prior to disclosure of personal health information, "the

researcher shall enter into an agreement with the custodian in which the researcher agrees to comply with the conditions and restrictions, if any, that the custodian imposes relating to the use, security, disclosure, return or disposal of the information.” I cannot agree with the University when it says that only “[e]ffective communication is the requirement,” referring to communication by the University. The *Act* is clear that there must be an agreement, which clearly requires that custodians must have agreed to these amendments and communicated such acceptance in order for them to be reflected in the research agreement.

[96] The University indicated that it only sent emails advising of changes to the Provider Agreement for the most significant changes: when it expanded the sharing of information with ICES and with the Diabetes Action Canada repository, and when it moved to full chart extraction. The University did not provide notification of other, more minor changes that it regarded as administrative or clerical in nature.

[97] I accept that the University reached out twice by email to notify custodians of the significant changes that it was making to its research plan. The issue is that the University set up a system in which it acted as though custodians had accepted the changes if it did not hear otherwise. I am not convinced that in these circumstances, silence in response to an email could amount to acceptance of the proposed amendments to the Provider Agreement relating to the increase in the scope of personal health information that UTOPIAN collected.

[98] First, with respect to the mode of communication, I note that emails can be inadvertently filtered to spam folders, where they are unlikely to be found or read. They can be overlooked altogether or put aside by busy practitioners to be looked at later, only to be forgotten about. Second, even though the emails were sent with read receipts, a read receipt does not indicate whether the email has been read. It is only able to indicate whether the email has been opened.

[99] Third, and most significantly, this was not some minor change. The collection of documents in picture-like formats greatly expanded the types of documents that UTOPIAN collected. As the University itself states, UTOPIAN began collecting “essentially all clinically relevant information contained within the EMR,” including direct patient identifiers,¹¹ noting that having “unstructured free text data also makes the data more amenable to artificial intelligence/text-mining technology.”

[100] Where the University did not receive questions or an opt-out, the University would begin collecting all clinically relevant information from those EMRs. For custodians who did not open or read the emails, this could occur without them knowing that they were contributing more personal health information. This approach chosen by the University

¹¹ Earlier versions of the Provider Agreement stated that direct patient identifiers would not be retained, but following the 2020 expansion of collection, UTOPIAN collected and retained these identifiers. Additional information regarding the changes between different versions of the Provider Agreement can be found at paragraph 142.

leaves open the possibility that disclosure could occur without those custodians agreeing to do so. It is not reasonable for the University to assume that non-response to an email indicates acceptance of the terms included within in the context of the significant changes that occurred in 2020, when the University cannot even be certain that the contents of the email were ever read.

[101] I conclude that the evidence of the University does not satisfy me that it effectively amended the Provider Agreement simply by sending the Amendment Email. In my view, the University should instead have taken steps to ensure that custodians clearly, unambiguously and unequivocally communicated their acceptance of the proposed amendment to the Provider Agreement, rather than relying on silence. In the context of this case, this could have been done by sending copies of the revised Provider Agreement with the updated consent form and requesting that custodians execute these, by including a link in the email for the custodians to click to indicate acceptance, or similar means.

[102] Given the significance of this change, both for custodians and for patients, I am not satisfied that providing an update by way of emails to busy practitioners and relying on silence was sufficient to amend the Provider Agreement executed by the custodians. To resolve this issue going forward, I recommend that the University ensure that it has a valid research agreement in place with each custodian, including any significant amendments thereto.

c. Did the University comply with the conditions, if any, specified by the research ethics board in respect of the research plan pursuant to section 44(6)(a) of the Act?

[103] Among the conditions set out in the REB-approved research plan were that the University take de-identification steps sufficient to de-link that information from identifying information and conduct site visits of custodians' offices to ensure that custodians were providing the required notice to their patients. The University's position is that it complied with both these conditions. I will address both of those obligations, and the University's compliance with them, in turn.

De-identification Processes

[104] In their letter to the IPC, the anonymous complainants raised particular concerns with the 2020 expansion in the scope of the UTOPIAN project to include unstructured free text fields, specifically mentioning the collection of picture-like documents and the possibility that UTOPIAN may transfer patients' entire medical chart to outside entities. The complaint noted that UTOPIAN keeps linkage files, allowing UTOPIAN to re-identify the personal health information in future. The complainants also raised concerns with possibility that identifying information may be leaked, or that other privacy breaches may occur.

[105] Taken together, it appears that the complainants are concerned about the

adequacy of the de-identification process and the disclosure of potentially identifying information from UTOPIAN to other parties. I therefore asked for detailed information about UTOPIAN's de-identification procedures, including whether the information available to researchers in the UTOPIAN database could identify individuals. Based on the concerns in the complaint, I also sought assurances from the University that researchers would not be provided with the picture-like documents or other identifying information.

[106] The University stated that the research ethics board's approval of UTOPIAN was conditional on "de-link[ing] identifying information," noting that "UTOPIAN must also de-link identifying information in accordance with the de-identification methods upon which its approvals are based." I agree that de-linking identifying information from the personal health information provided prior to including that information in the research database is a condition of the research plan specified by the REB.

[107] UTOPIAN retained the ability to re-identify the information for the purposes of recruitment for interventional studies that have their own project-specific REB approvals. However, UTOPIAN asserts in its research plan that "[no] data that can identify patients or their primary care providers will be part of the research database." The research plan set out the steps for UTOPIAN to take to ensure that it did not provide researchers with personal health information.

[108] To begin, UTOPIAN assigned a unique system-generated EMR ID number to each patient within a participating practice's EMR. Then, UTOPIAN used this number to randomly generate another number, the UTOPIAN ID, which was used as a patient study number. These numbers were saved together in an EMR Data Linkage File.

[109] The University states that the EMR Data Linkage File allowed it to recognize if data from an upload belongs to a patient whose data is already in UTOPIAN. In addition, in cases where a patient withdraws from the project, this data linkage file allows UTOPIAN to identify and remove that patient's personal health information from the database. The University states that one cannot re-identify a patient from these two numbers unless the EMR Data Linkage File is linked back to the source EMR data, and neither the source data nor the EMR Data Linkage File were available to researchers.

[110] The University also described the process by which it de-linked the identifying information it collected from custodians. The research plan notes that UTOPIAN extracted a wide range of health data – "conditions, encounters, prescriptions, referrals, laboratory tests, diagnostic tests, consult and referral letters, immunization data, essentially all clinically relevant information contained within the EMR." This data was either entered directly into the EMR's fields or captured from free-text fields and in picture-like documents.

[111] UTOPIAN extracted health data from the various fields within the EMRs and scanned it for identifying numbers belonging to the patient (driver's license, health card, etc.), addresses and other contact information, such as names, surnames addresses,

postal codes, etc. It removed these identifiers and replaced them with “proxies” – essentially, randomly assigned placeholders. UTOPIAN then moved this data (with the proxy values, rather than the original identifiers) into a staging area, separate from where the identifying information was stored.

[112] A similar process occurred for the picture-like documents, but with an added step. Because these are images, UTOPIAN could not just pull the data directly, as it did with data that is entered into the EMR fields. Instead, it first had to process these documents to locate the relevant information within them. It did so using an optical character recognition system. Once located, any identifying information was likewise replaced by proxies before moving to the next step in the process. The picture-like documents containing identifying information remained in the secure area, apart from the database.

[113] The proxy values assigned were kept constant throughout sections of a medical record, but not throughout the entire record. The University provided an example of one proxy name replacing the actual name in a medication section of a record, but a different name being used as the proxy in the allergy section of that record.

[114] The indexes linking the values and proxy values were stored in a secure zone with other personally identifiable data. They were only accessible by a limited number of authorized UTOPIAN personnel, and not by researchers.

[115] The University stated that these proxy values were generated randomly, and not based on a hashing algorithm. Hashing algorithms take one input and transform it into another, the “hash.” When using a hashing algorithm, the same input always results in the same hash. An individual who can see the hashes and knows that a hashing function was used will be able to infer that the inputs were the same value, even if they do not know what that value is. Using random number generation for proxies means that the same inputs can have different proxy values, providing an extra layer of protection.

[116] Once the proxy values were substituted for the identifying information, the University states that UTOPIAN performed a final manual check on a sample of the data. This was to ensure that the de-identification process was executed properly, and no identifying information remains. When this was completed, the data was then available for research.

[117] Researchers do not have direct access to this database. Instead, UTOPIAN staff compile sets of data relevant to the researcher’s project and make these datasets available to researchers within the SAVE environment. Prior to doing so, the datasets are further reviewed by a SAVE systems administrator to confirm that the de-identification was completed effectively. This administrator also removes any data in cell sizes of five or less. As part of this process, new proxy values are generated for each research dataset, such that different researchers see different proxy values for the same underlying value.

[118] Researchers can then work with the de-identified dataset within SAVE but cannot

modify or export this data. Furthermore, researchers are required to enter into the Researcher NDA with the University. Under this agreement, the researcher stipulates that they will not attempt to learn the identity of any individuals associated with information that is “non-public, confidential or proprietary in nature” that it accessed via UTOPIAN, and agrees not to download any data.

[119] The Researcher NDA and the limitations on researchers’ ability to manipulate the data reduce the risk of researchers linking UTOPIAN datasets with those they may have access to from outside sources. If researchers were able to link datasets, they may be able to assemble a greater collection of data about any one individual and, depending on the nature and extent of the information they had about that individual, potentially identify them.

[120] When asked, the University stated that it had not conducted a specific re-identification risk study for UTOPIAN, but “assesses the risk of re-identification as low.”

[121] The complainant flagged concerns about documents containing identifying information being provided to third parties. From the University’s description of the functioning of the UTOPIAN database, identifying information can only be accessed by authorized UTOPIAN personnel. I have not found any evidence to the contrary.

[122] The complainant’s assertion that UTOPIAN keeps linkage fields, allowing for re-identification, is correct. However, the purposes and conditions for retaining this linkage file are specifically set out in UTOPIAN’s research plan that was approved the REB. From the information provided, it appears that UTOPIAN has procedures in place to comply with the conditions specified by the REB¹².

[123] However, based on the highly sensitive nature of the personal health information that the University obtained from providers, a re-identification study would be an appropriate step to take. It would provide reassurance to the public and custodians, and could flag whether any additional procedures would assist in further mitigating risks of potential re-identification. UTOPIAN is no longer operating but the University stated that it intends to submit a new application to the REB for the continued use of archived UTOPIAN database for research purposes.

[124] I recommend that the University conduct a re-identification study of the UTOPIAN database for that purpose. This study should be done in accordance with the best practices set out in IPC’s *De-identification Guidelines for Structured Data*¹³ and ISO/IEC 27559:2022, and should include an analysis of a specific dataset.

¹² I make no determination as to whether the personal health information has been de-identified, as a determination of that nature would likely require a re-identification study, such as I recommend below. As part of my review, I have examined whether the University is complying with the obligations in its research plan relating to the de-linking of personal health information.

¹³ <https://www.ipc.on.ca/resource/de-identification-guidelines-for-structured-data/>

Site Visits

[125] Based on the first research plan, the UTOPIAN Practice Coordinator was required to “visit each practice site regularly to monitor sites’ compliance with posters being displayed.” This obligation continued through the life of the project. In the research plan, including in the section addressing why UTOPIAN should not require patient consent, the University cites the Patient Information Posters as a means by which patients are informed of their right to withdraw their personal health information from the project. The site visits ensure that the most recent version of the Patient Information Poster is still in place and visible and are themselves a condition of the research plan for the purposes of s. 44(6)(a). Accordingly, I asked the University about the frequency of such site visits.

[126] In response, the University stated that though it did not keep records of site visits. It conducted such visits prior to the pandemic but paused these during the pandemic out of safety concerns. The University’s position is that the requirement that visits be “regular” should be read pragmatically, to allow for a pause caused by the pandemic.

[127] Though I understand why site visits would not have been possible for health and safety reasons in the earlier days of the COVID-19 pandemic, that is no longer the case, and has not been for some time. The University could have, and should have, resumed these site visits at some time in the intervening years.

[128] While “regularly” is not defined in the research plan, pausing site visits since early 2020 cannot meet any standard of regular visits. I find that the University’s cessation of site visits entirely since early 2020 is a failure to comply with a condition of its research plan specified by the REB, contrary to section 44(6)(a) of the *Act*.

[129] More importantly, the University asserts that the shift to virtual care since the pandemic rendered these visits “of very limited value” and stated that the site visit requirement should have been amended, though it was not. While I understand the University’s reasoning, carrying this logic further brings into question the adequacy of the mechanism by which patients are to be notified. Given that the posters themselves are unlikely to be seen by patient’s attending virtual care visits, patients – especially those new to the practice – are unlikely to be aware that their personal health information was being used for the UTOPIAN project. If the University assessed that the posters were no longer providing adequate notice to patients of their right to opt-out of the project given the move to virtual visits, the proper response is not to cease the required site visits. Rather, the university should have varied the research plan in such a way as to propose an alternative and more effective form of notice during a time when virtual care was more prevalent. In its new application to the REB for the use of the archived UTOPIAN database, I recommend that the University update the means of notifying patients other than only by way of notices in physical offices in a context where virtual visits have become more prevalent since the pandemic.

d. Did the University use the personal health information provided only for the purposes set out in the research plan as approved by the research ethics board pursuant to section 44(6)(b) of the Act?

[130] The complaint expresses concerns about UTOPIAN uploading personal health information that may be “intentionally taken, transferred, used, altered, stored, and sold.” In addition to this, the complaint later states that UTOPIAN “offers to sell the appropriated data to other organizations, research teams and individual researchers.”

[131] When asked about the allegations relating to the sale of data, the University stated:

At no time did UTOPIAN, nor will POPLAR, buy or sell data. The University has charged individual researchers who access the data resource fees for certain services on a cost-recovery basis. These fees are strictly for services provided in connection with an approved research study (e.g., fees for project facilitation, methodological support, data management and extraction, data analysis, administrative fees). [Emphasis in original]

[132] I have reviewed the University’s description of the operation of the UTOPIAN database. I am satisfied that nothing in the materials provided indicates that the University sold personal health information to third parties. UTOPIAN disclosed clinical datasets to prescribed entities and similar bodies (such as ICES¹⁴) and, separately, provided researchers with the use of de-identified datasets on a fee for service basis. Both of these activities are addressed in the research plan, as approved by the REB, and the Provider Agreement.

[133] I find that in its operation of UTOPIAN, the University used the personal health information provided only for the purposes set out in the research plan as approved by the REB and that no sale of personal health information occurred.

[134] While I am satisfied that there was no selling of personal information, it is worth considering the reason that the complainants may have raised this issue to our office. The complainants state that they have tried to raise their larger concerns with stakeholders but have “essentially been shut down.” I am not aware of the extent or content of any discussions. I do note that greater transparency could potentially have

¹⁴ I also note that section 17 of Regulation 329/04 to the *Act* permits disclosure of personal health information to prescribed entities in certain circumstances, reading as follows:

Despite clause 44 (6) (d) of the Act, a researcher may disclose the information to an entity prescribed under subsection 45 (1) of the Act, to a person prescribed for the purposes of clause 39 (1) (c) of the Act for use in a registry compiled or maintained by that person, or to another researcher if,

(a) the disclosure is part of a research plan approved under section 44 of the Act; or
(b) the disclosure is necessary for the purpose of verifying or validating the information or the research.

averted the complaint.

[135] That custodians who filed the complaint reached the conclusions that they did based on the limited information they had and felt they could not address their concerns directly with the UTOPIAN researchers, raises some concern regarding transparency. This may also be partly the result of the University's choosing to provide custodians with the Provider Agreements, instead of the research plan and associated REB approvals. While I accept that the University was attempting to provide information to custodians in what they believed to be the most effective way, this abbreviated format may have contributed to the suspicion and distrust on the part of at least some of the custodians, such as those who made this complaint. In future projects, I recommend that the University exercise greater transparency with any contributing custodian and open lines of communication to instill a sense of trust among them and their patients.

e. Did the University provide notice to custodians of a breach of section 44(6) of the Act, or a breach of the research agreement, pursuant to section 44(6)(f) of the Act?

[136] Section 44(6) requires that researchers provide custodians with notice of any contravention of the provisions of subsection 44(6), or of the agreement between the parties under subsection 44(5). The University outlined their notice obligation in the Provider Agreement, which states as follows:

The research team agrees to report to you any breach of confidentiality and/or security respecting your patients' health data where such breach may result in the identification of one or more of your patients. Regardless of whether or not the data involved in the breach results in identification of your patient or patients, UTOPIAN, and/or CPCSSN, as the case may be, will take immediate and reasonably necessary steps to both remedy the breach and prevent similar occurrences in the future. You will be notified as soon as reasonably possible if such breach involves one or more of your patients' data.

[137] The University provided notice of the breach to the custodians whose data was uploaded on January 5, 2023. The University stated that "this inadvertent upload of personal health information is considered to be inconsistent with PHIPA", and that they were providing notice "per our legal obligations." That notice letter outlined that the contravention had been contained, and that the uploaded data would be destroyed. It also informed custodians that they needed to provide notice to patients who may have been affected and asked that custodians share the letter with their EMR provider.

[138] Regarding the 2018 uploads during the lapse of the REB approval, the University states that it did not provide notice to the affected custodians, stating that this was "due to oversight." The University is required under section 44(6)(f) to provide notice in writing if it "becomes aware of any breach" of the research agreement. The University stated

that it became aware of the expiry of the REB approval before it submitted its October 5, 2018 application to the REB. It should have provided notice of this breach to custodians at that time.

[139] I find that the University provided notice of the 2023 breach of section 44(6) as required by section 44(6)(f) of the *Act*, but did not do so in respect of the 2018 unauthorized collections.

f. Did the University comply with the research agreement as required under section 44(6)(g) of the Act?

[140] The University has stated that for the vast majority of cases, the Provider Agreement is the research agreement between UTOPIAN and custodians for the purpose of s. 44(5) of the *Act*. Some small number of custodians entered into individualized research agreements with the University; as noted above, I am not addressing those agreements in this Decision.

[141] As discussed above, the University purported to amend the Provider Agreement by informing custodians of significant changes, and treating the silence of custodians as acceptance of these changes that amended the Provider Agreement. I do not agree that the non-response of custodians was sufficient to amend the Provider Agreement, so the Provider Agreement that governs the relationship between the custodian and the University is the version executed by the custodian. Therefore, the applicable version of the Provider Agreement will vary between custodians, depending on when they signed on to UTOPIAN and executed their Provider Agreement.

[142] The University has made changes to UTOPIAN since its genesis, and the different versions of the Provider Agreement reflect these changes. The most significant difference between the versions is the description of the personal health information collected by UTOPIAN which changed in 2020 to reflect the increased scope of personal health information collected. While all versions of the Provider Agreement specify categories of information that UTOPIAN collected, the pre-2020 Provider Agreements limit any additional collection to "other data related to chronic conditions" and note that no patient direct identifiers are retained. In contrast, the later Provider Agreements clearly expanded the collection of personal health information to include patient direct identifiers, information from picture-like documents, and "essentially all clinically relevant information contained within the EMR."

[143] There is no dispute that since 2020, UTOPIAN collected, used, and retained more personal health information that it did previously. This includes collection, use, and retention of significant personal health information (derived from picture-like documents) and patient direct identifiers.

[144] If a custodian's executed Provider Agreement does not include this increase in scope (as would be the case for custodians who signed on prior to these 2020 changes),

then collection of these additional data elements is not included within the activities agreed to in the Provider Agreement. In such cases, the University is not in compliance with the applicable research agreement as required by section 44(6)(g) of the *Act* in so far as the University collected, used, and retained personal health information not set out in the applicable Provider Agreement.¹⁵

Conclusions regarding the University's compliance with s. 44

[145] In summary, I find that the University failed to comply with the section 44 obligations for researchers in the following ways:

- The University failed to provide custodians with a research plan and the REB decision approving that research plan, contrary to s. 44(1)(a).
- Between August 7, 2018 and October 4, 2018, and again on January 5, 2023, the University collected personal health information without a valid REB approval in place contrary to section 44(1)(a).
- The University failed to conduct the site visits set out in the research plan from early 2020 onwards, which is a failure to comply with a condition specified by the REB in respect of the research plan, contrary to section 44(6)(a).
- The University did not provide notice of the 2018 unauthorized collections of personal health information as required under the Provider Agreement, contrary to section 44(6)(f).
- In instances in which the University collected, used, and retained personal health information beyond what was specified in the applicable research agreement, the University did not comply with section 44(6)(g).

[146] The University has paused all activities related to the UTOPIAN database. However, the University has also stated that it is currently developing an application to the REB to permit the continued use of the archived UTOPIAN database for research purposes. I recommend that the University:

1. Update the means of notifying patients other than only by way of notices in physical offices, in a context in which virtual care visits are more prevalent.
2. Conduct a re-identification assessment in accordance with the best practices set out in IPC's *De-identification Guidelines for Structured Data* and ISO/IEC 27559:2022, including analysis of a specific dataset, to test the robustness of its de-identification procedures.

¹⁵ This conclusion is subject to the caveat noted above at footnote 10.

3. Ensure that it has a valid research agreement in place with each custodian, including any significant amendments thereto, and that it is in compliance with this agreement.
4. Exercise greater transparency generally with contributing custodians and open lines of communication to instill a sense of trust among them and their patients.

2. Is the creation of the UTOPIAN database "research" pursuant to sections and 44 of the Act?

[147] The above analysis was done on the premise that UTOPIAN's activities were research under the *Act*, without deciding that point. This allowed me to consider whether, if that were the case, the University met the conditions imposed on researchers, as set out in section 44 of the *Act*. However, the complaint sent to this office questioned whether the creation and operation of the UTOPIAN database was even "research" as contemplated under the *Act*.

[148] Section 2 of the *Act* defines research as "a systematic investigation designed to develop or establish principles, facts or generalizable knowledge, or any combination of them, and includes the development, testing and evaluation of research." This definition is silent on the duration and nature of individual research projects. Section 44 similarly does not address these specifics.

[149] The University describes UTOPIAN as a "critical initiative that [supports] primary care research in Ontario." The University states that research using EMR data is "crucial to understanding the management of acute and chronic disease, health promotion and disease prevention in the community." UTOPIAN is no doubt a significant endeavour, and to carry it out, the University has set up an extensive database created from the personal health information of nearly 600,000 Ontarians.

[150] To reiterate, the original purposes listed by the University when it applied to the REB for approval were:

- To create a database with anonymized patient data from EMRs of primary health care providers
- To provide accessible data options for research and public health surveillance
- To devise algorithms or other processes to enable automated EMR data collection, data de-identification, and other data processes.

[151] The language of section 44 of the *Act* stipulates what steps researchers must take for health information custodians to disclose personal health information for research purposes without individual consent. Its provisions clearly map on to custodians disclosing personal health information to carry out discrete, time-limited research projects. The University takes the position that this section also permits the disclosure of personal

health information to create a database that will, in turn, be used for the purposes of ongoing and future research projects.

[152] The University has cited the importance of using EMR data for research into various matters related to health and disease. I do not disagree. However, UTOPIAN and its fellow PBLRNs are not the only avenue for large-scale research and/or analysis of this nature to take place within the bounds of *PHIPA*. Section 45 of the *Act* was specifically designed to allow for the creation of prescribed entities to whom health information custodians may disclose personal health information without consent, which in turn can be made available to others for research purposes, as contemplated by section 18(4) of Reg. 329/04. Indeed, one of the prescribed entities, ICES, has an existing data linkage with UTOPIAN and offers similar research services to UTOPIAN.

[153] Prescribed entities under section 45 of *PHIPA* are large-scale data repositories that operate as an ongoing concern. They are afforded greater flexibility under the *Act* to collect personal health information from custodians without patient consent for the purpose of analysis or compiling statistical information with respect to “the management of, evaluation or monitoring of, the allocation of resources to or planning for all or part of the health system, including the delivery of services”. However, with this greater flexibility comes greater accountability in the form of high levels of protection and oversight by our office.

[154] Specifically, prescribed entities must have in place practices and procedures to protect the privacy of the individuals whose personal health information it receives and to maintain the confidentiality of the information. Such practices and procedures must be approved by the IPC every three years in order for it to continue operating.

[155] The University has characterized UTOPIAN as a research project under section 44 of the *Act*. However, using section 44 to authorize large-scale research platforms that operate as an ongoing concern, such as UTOPIAN, can lead to many practical difficulties given the awkward fit. Difficulties arise because it can be unclear how to apply the section 44 requirements to general research platforms that, unlike individual research projects, have no narrow research purpose or end date. Keeping custodians regularly up to date with the evolving research plan and multiple REB approvals, and securing their agreement to continue to participate on an ongoing basis, can pose significant practical challenges, as we have seen materialize in this case.

[156] I recognize that operating under section 45 carries a high level of oversight regarding privacy and confidentiality practices and procedures. However, it also provides the greater flexibility that platforms like UTOPIAN would benefit from to operate as an ongoing concern and provides custodians and patients with the necessary assurances that the platform is being regularly reviewed by our office to ensure it is operating securely and in compliance with the law. The clear rules and guidance for s. 45 entities, when compared with the efforts required to comply with s. 44 requirements, could well be to the University’s advantage, as well as to the advantage of Ontarians, because their

personal health information is more consistently protected while getting the benefit of health system research and analysis.

[157] Given that UTOPIAN is no longer collecting personal health information pursuant to section 44 of the *Act*, it is not necessary for me to determine whether the creation of the UTOPIAN database is "research" pursuant to the *Act*. However, the distinctions outlined above between specific research projects operating under section 44, and similar activities occurring within prescribed entities under section 45, is an important one for bodies contemplating research and health sector analysis to consider when designing any future such endeavors.

3. Should this matter proceed to adjudication at the IPC, where a potential order may be issued?

[158] As discussed above, I have found several areas in which the University did not comply with s. 44 of the *Act*.

[159] However, as of February 2024, all UTOPIAN activities have been paused. The University has informed this office that it will be submitting a new application to its REB for the continued use of the archived UTOPIAN database for research purposes, but that it is still in the process of developing that application. While that research project is not currently before this office, this Decision provides important findings and recommendations that the University should consider in that process.

[160] Also, given the time that has elapsed since the 2018 unauthorized collections of personal health information, it is not necessary for this matter to proceed to adjudication for the purposes of ordering the University to provide notice of an unauthorized collection of personal health information pursuant to section 44(6)(f).

[161] Based on the current state of affairs, I do not see any utility or necessity in having this file proceed to the adjudication stage in this matter.

[162] Therefore, in accordance with my delegated authority under the *Act*, and for the reasons set out above, this review will be concluded without proceeding to the adjudication stage.

CONCLUSIONS:

1. Assuming UTOPIAN constitutes research as contemplated by section 44 of *PHIPA*, I find that the University failed to comply with the section 44 obligations for researchers in the following ways:
 - a. The University failed to provide custodians with a research plan and the REB decision approving that research plan, contrary to s. 44(1)(a).

- b. Between August 7, 2018 and October 4, 2018, and again on January 5, 2023, the University collected personal health information without a valid REB approval in place contrary to section 44(1)(a).
 - c. The University failed to conduct the site visits set out in the research plan from early 2020 onwards, which is a failure to comply with a condition specified by the REB in respect of the research plan, contrary to section 44(6)(a).
 - d. The University did not provide notice of the 2018 unauthorized collections of personal health information as required under the Provider Agreement, contrary to section 44(6)(f).
 - e. In instances in which the University collected, used, and retained personal health information beyond what was specified in the applicable research agreement, the University did not comply with section 44(6)(g).
2. Given that all activities of UTOPIAN have been paused, it is not necessary for me to determine whether the creation of UTOPIAN in its previous form was “research” pursuant to the *Act*.
 3. This review will be concluded without proceeding to the adjudication stage.

RECOMMENDATIONS:

I recommend that the University:

1. Update the means of notifying patients of UTOPIAN other than only by way of notices in physical offices, in a context in which virtual care visits are more prevalent.
2. Conduct a re-identification assessment in accordance with the best practices set out in IPC's *De-identification Guidelines for Structured Data* and ISO/IEC 27559:2022, including analysis of a specific dataset, to test the robustness of its de-identification procedures.
3. Ensure that it has a valid research agreement in place with each custodian, including any significant amendments thereto, and that it is in compliance with this agreement.
4. Exercise greater transparency generally with contributing custodians and open lines of communication to instill a sense of trust among them and their patients.

Within six months of receiving this Decision, the University should report back to this office regarding the implementation of these recommendations.

Original Signed by: _____
Jennifer Olijnyk
Investigator

_____ May 21, 2024

POSTSCRIPT

I understand that Queen's has applied for, and been granted, approval to operate POPLAR by Queen's REB, though that project is not yet operational. During the investigation of this complaint, Queen's contacted the IPC to express its full cooperation in the IPC's ongoing review of this file, as it relates to POPLAR. They also provided me with POPLAR's research plan and the decision of the Queen's REB approving that plan. I commend Queen's for being proactive in getting in touch with this office regarding this matter.

I am providing a copy of this decision to Queen's University for their consideration. I ask that they consider my reasoning in this decision, and in particular, my comments as set out in paragraph 156 addressing how a research project in the nature of a large-scale database operating as an ongoing concern, such as this, may be ill-suited to comply with the requirements of s. 44, and whether it should operate under the requirements set out in s. 44 or s. 45 of *PHIPA*. The IPC, through its health policy department, remains available for further consultation on this matter under section 66(d) of the *Act*, which grants this office the discretion to offer comments on the custodian's actual or proposed information practices, on request.

APPENDIX 1

Health information custodian

3 (1) In this Act,

“health information custodian”, subject to subsections (3) to (11), means a person or organization described in one of the following paragraphs who has custody or control of personal health information as a result of or in connection with performing the person’s or organization’s powers or duties or the work described in the paragraph, if any:

1. A health care practitioner or a person who operates a group practice of health care practitioners.

2. REPEALED: 2020, c. 13, Sched. 3, s. 8 (3).

3. A health service provider or person or entity that is part of an Ontario Health Team and that provides a home and community care service pursuant to funding under section 21 of the *Connecting Care Act, 2019*, including a person or entity from whom the provider or Team has purchased the home and community care service.

4. A person who operates one of the following facilities, programs or services:

i. A hospital within the meaning of the *Public Hospitals Act*, a private hospital within the meaning of the *Private Hospitals Act*, a psychiatric facility within the meaning of the *Mental Health Act* or an integrated community health services centre within the meaning of the *Integrated Community Health Services Centres Act, 2023*.

ii. A long-term care home within the meaning of the *Fixing Long-Term Care Act, 2021*, a placement co-ordinator described in subsection 47 (1) of that Act, or a care home within the meaning of the *Residential Tenancies Act, 2006*.

ii. 1 a retirement home within the meaning of the *Retirement Homes Act, 2010*.

iii. A pharmacy within the meaning of the *Drug and Pharmacies Regulation Act*.

iv. A laboratory or a specimen collection centre as defined in section 5 of the *Laboratory and Specimen Collection Centre Licensing Act*.

- v. An ambulance service within the meaning of the *Ambulance Act*.
 - vi. A home for special care within the meaning of the *Homes for Special Care Act*.
 - vii. A centre, program or service for community health or mental health whose primary purpose is the provision of health care.
5. An evaluator within the meaning of the *Health Care Consent Act, 1996* or an assessor within the meaning of the *Substitute Decisions Act, 1992*.
 6. A medical officer of health of a board of health within the meaning of the *Health Protection and Promotion Act*.
 7. The Minister, together with the Ministry of the Minister if the context so requires.
 8. Any other person prescribed as a health information custodian if the person has custody or control of personal health information as a result of or in connection with performing prescribed powers, duties or work or any prescribed class of such persons.

Personal health information

4 (1) In this Act,

“personal health information”, subject to subsections (3) and (4), means identifying information about an individual in oral or recorded form, if the information,

(a) relates to the physical or mental health of the individual, including information that consists of the health history of the individual’s family,

(b) relates to the providing of health care to the individual, including the identification of a person as a provider of health care to the individual,

(c) REPEALED: 2020, c. 13, Sched. 3, s. 8 (7).

(c.1) is a plan that sets out the home and community care services for the individual to be provided by a health service provider or Ontario Health Team pursuant to funding under section 21 of the *Connecting Care Act, 2019*,

(d) relates to payments or eligibility for health care, or eligibility for coverage for health care, in respect of the individual,

(e) relates to the donation by the individual of any body part or bodily substance of the individual or is derived from the testing or examination of any such body part or bodily substance,

(f) is the individual's health number, or

(g) identifies an individual's substitute decision-maker.