



Article

Implementation of Gender Identity and Assigned Sex at Birth Data Collection in Electronic Health Records: Where Are We Now?

Hale M. Thompson ^{1,*}, Clair A. Kronk ², Ketzal Feasley ¹, Paul Pachwicewicz ¹ and Niranjan S. Karnik ¹

¹ Department of Psychiatry and Behavioral Science, Rush University Medical Center, Chicago, IL 60612, USA; ketzel_feasley@rush.edu (K.F.); paul_pachwicewicz@rush.edu (P.P.); niranjan_karnik@rush.edu (N.S.K.)

² Department of Biomedical Informatics, University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA; kronkcj@mail.uc.edu

* Correspondence: hale_thompson@rush.edu

Abstract: In 2015, the United States Department of Health and Human Services instantiated rules mandating the inclusion of sexual orientation and gender identity (SO/GI) data fields for systems certified under Stage 3 of the Meaningful Use of Electronic Health Records (EHR) program. To date, no published assessments have benchmarked implementation penetration and data quality. To establish a benchmark for a U.S. health system collection of gender identity and sex assigned at birth, we analyzed one urban academic health center's EHR data; specifically, the records of patients with unplanned hospital admissions during 2020 (N = 49,314). Approximately one-quarter of patient records included gender identity data, and one percent of them indicated a transgender or nonbinary (TGNB) status. Data quality checks suggested limited provider literacy around gender identity as well as limited provider and patient comfort levels with gender identity disclosures. Improvements are needed in both provider and patient literacy and comfort around gender identity in clinical settings. To include TGNB populations in informatics-based research, additional novel approaches, such as natural language processing, may be needed for more comprehensive and representative TGNB cohort discovery. Community and stakeholder engagement around gender identity data collection and health research will likely improve these implementation efforts.

Keywords: SO/GI data; nonbinary; population health; clinical informatics; transgender health disparities



Citation: Thompson, H.M.; Kronk, C.A.; Feasley, K.; Pachwicewicz, P.; Karnik, N.S. Implementation of Gender Identity and Assigned Sex at Birth Data Collection in Electronic Health Records: Where Are We Now? *Int. J. Environ. Res. Public Health* **2021**, *18*, 6599. <https://doi.org/10.3390/ijerph18126599>

Academic Editors:
Cristiano Scandurra and
Kimberly Balsam

Received: 30 May 2021
Accepted: 16 June 2021
Published: 19 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Who Counts as Transgender? What Counts as Sexual Orientation? The Case of Patient Doe

In a 2020 quality improvement report to determine the prevalence of hospital patient electronic health records (EHR) that include gender identity (GI) and sexual orientation (SO) data, 41 patients had a sexual orientation listed as “transgender” ($n = 13$) or “neither exclusively man nor woman” ($n = 28$). A quality check of one such patient's record, patient “Doe”, indicated a possible mapping error of a gender identity noted by a provider in the social history, instead of the sexual orientation that was listed in the demographics table. Although the report indicated a “transgender/male-to-female” sexual orientation, patient Doe's demographic SO/GI tab read: male (assigned sex at birth), male (gender identity), heterosexual (sexual orientation), and had no pronouns indicated. Doe is referred to with he/him pronouns throughout his record. In the social history section, however, where the institution had been collecting SO/GI data prior to 2020, a note reads that patient Doe had seen a physician about “transgender reassignment, male-to-female.” Doe's photo in the patient snapshot appears very typically “male” in presentation, and the photo of Doe's driver's license shows “sex: M.” Notably, the surgical history section dating back years, reads “transgender s/p M to F reassignment surgery,” and medication notes through 2021 indicate the use of estrogen.

1.2. Background and Significance

Little is known about the quality of electronic health record data regarding gender identity (GI) and sexual orientation (SO). From Doe's case, we can deduce that the "transgender" sexual orientation reflects a mapping error since his actual record's SO data field indicates heterosexual. But how should one count this patient in terms of gender identity? In our quality improvement analysis that follows, patient Doe is counted as cisgender, and it appears that is how he wants to be recognized in the clinical context. Doe's structured data fields for legal sex, assigned sex at birth (ASAB), and GI are aligned as male, but keywords in the social history and clinical notes of the EHR, indicate the presence of gender-confirming surgical and hormonal interventions from an outside provider, leading us to interpret patient Doe as a transgender person who has medically transitioned but has not legally or entirely socially transitioned. EHRs have purposes that extend beyond the patient in the clinical context. Increasingly, EHR data are the basis for a range of health research, and in the context of research and equity, perhaps we would want to count transgender patients, such as Doe, who are not "out" in clinical settings where they may not feel comfortable or safe. This case provides insight into the uses, misuses, and the importance of SO/GI EHR data collection while also extending the question of who counts as transgender and, increasingly, nonbinary [1–5].

Formulations of gender identities (see for example, [6–8]) and modalities [9] beyond U.S. borders have been contested and redefined over time at local levels across the globe. International organizations like the United Nations (UNESCO), World Professional Association of Transgender Health (WPATH), and the World Health Organization's (WHO) International Classification of Diseases (ICD) have made more sweeping attempts at the establishment of global definitions moving from, for example, an individual pathology (i.e., Gender Identity Disorder in ICD-9) to a more socially circumscribed illness (i.e., Gender Dysphoria in ICD-10) to a relatively neutral status (i.e., Gender Incongruence coming in ICD-11) [10–12]. Both WPATH and UNESCO have established more flexible definitions for transgender and nonbinary (TGNB) statuses, essentially establishing TGNB as any gender that differs from the one associated with one's sex assigned at birth [13]. This broad definition does not negate the complexities of gender across and within nations, cultures, histories, and healthcare systems, and reflects a growing prevalence and/or awareness of nonbinary identities [1,14–16].

In the United States, when insurance coverage for gender-affirming care expanded with the Affordable Care Act (ACA), the Health Resources and Services Administration (HRSA), the Joint Commission, and the Institute of Medicine recommended routine collection of gender identity (GI) in healthcare settings in order to improve access to and quality of care. In 2014, the ACA's Meaningful Use Stage 3 required Medicaid provider EHRs to have capacity to collect gender identity in a structured data field. By 2019, 53% ($n = 360$) of the 680 U.S. hospitals that participated in the Human Rights Campaign 2019 Health Equality Index reported collecting GI data, and 87% ($n = 313$) of those reported using the two-step method [17], capturing both GI and ASAB. Rush University Medical Center (RUMC), Rush Oak Park Hospital, and Rush Copley Medical Center in suburban Aurora are three of those 360 U.S. hospitals.

A 2019 study of 1367 U.S. Health Centers reporting to HRSA's 2016 Uniform Data System analyzed the results to date and found that GI data was reported missing for 68% of patients [18]. This rate of uptake indicates major barriers to achieve the benefits ascribed to GI data collection and supports the contention that 40 years of insurance exclusions created a vast clinical training void for which GI data fields have not yet been able to compensate [19]. That is, patients and providers alike may hesitate to complete these fields as users mostly do not know how to interpret or communicate the information effectively [20–23]. The miscommunication or misuse of GI data in a clinical setting, due to a lack of provider training, may be a patient safety issue [24]. Another patient-centered reason for low penetration may have to do with the appropriateness of GI data collection during a particular encounter [25,26]. At the same time, secondary uses of clinical

informatics and EHR data are the core of learning health systems science and offer a new, pragmatic research pathway for trans and nonbinary population health research [27].

1.3. Objective

The primary objective of this analysis is to assess markers of implementation effectiveness: the penetration and quality of GI-related data collection and, at the same time, to establish a 2020 benchmark for collection of gender identity, sex assigned at birth, as well as sexual orientation at this academic health center and for health systems across the United States. Secondly, this analysis may provide insights into implementation barriers and facilitators as well as direction for gender identity, assigned sex at birth, and sexual orientation data quality standards.

2. Materials and Methods

2.1. Setting and Sample

Rush University Medical Center is a 727-bed hospital, tertiary care academic health center that serves Chicago and is located in the Near West Side. The Human Rights Campaign's Health Equality Index has designated RUMC a national leader in the care of LGBTQ patients for 12 consecutive years [17]. RUMC uses Epic (Epic Systems Corporation, Verona, Wisconsin) for its electronic health record system. As with all EHR systems, Epic captures "sex" and refers to the sex associated with a patient's insurance card or driver's license. In addition, from approximately 2013 through 2019, RUMC captured SO/GI data in the social history section of the EHR. The social history section captures data related to the social determinants of health and structural vulnerabilities [28] that impact patient health outcomes (e.g., patient status with regard to housing instability, the legal system, intimate partner violence, substance misuse, sex work, etc.). This domain of the EHR is not accessible to the patient, is not considered essential medical information, and is therefore also easily bypassed or overlooked by providers.

At the beginning of 2020, RUMC deployed the Epic sexual orientation (SO) and gender identity (GI) SmartForm, moving SO/GI data capture from social history to the demographics section of the EHR across the system. The demographics section has fields for ASAB, GI, and SO (see Figure 1); these are three of the SO/GI fields added to Epic in accordance with the 2015 Stage 3 Meaningful Use rules. Shifting the SO/GI data fields from social history to demographics enables a RUMC patient to access those fields and add or manually change their ASAB, GI, and/or SO once they are registered as a patient in the system. Although the legal name and legal sex appear in this tab, neither the patient nor provider is able to enter or change data in those fields, which are populated and affixed when a patient chart is established based on official insurance or identification. During an unplanned admission, the EHR workflow prompts the intake provider to complete these fields during intake or initial interview with a patient. However, the EHR does not have hard stops in this SmartForm, and the provider may bypass the fields due to time constraints, patient accessibility/unconsciousness, or discomfort with the questions. If the patient was already established as a patient from prior visits or admissions, the SO/GI fields may have been completed by providers known for gender-affirming care, or the patient.

Sexual Orientation and Gender Identity SmartForm

Inform the patient that anything entered here will be visible to anyone with access to this legal medical record.

Sexuality

Patient's sexual orientation:

Legal Information

Legal first name:

Legal last name:

Legal sex:

Gender Identity

Autofill with default responses for:

Patient's gender identity:

Patient's sex assigned at birth:

Patient pronouns:

© 2021 Epic Systems Corporation

Figure 1. Epic Demographic Sexual Orientation (SO) and Gender Identity (GI) SmartForm (© 2021 Epic Systems Corporation) implemented for SO/GI data collection at Rush University Medical Center in 2020.

The inclusion criteria for the quality improvement project were all unplanned adolescent and adult hospital patients (≥ 12 years of age) admitted between 1 January 2020 and 31 December 2020 ($N = 49,314$). This cohort includes all patients seen in the Emergency Department or admitted to the hospital, typically by way of the Emergency Department. The acute care setting and unplanned admissions are typically unrelated to planned, gender-affirming procedures with providers who are known for gender-affirming care. It is critical to assess the extent to which GI and ASAB data are collected in an acute care setting in order to evaluate patient experience, provider competencies, and to identify comorbidities and health outcomes associated with trans or nonbinary patients unrelated to gender-affirming procedures.

2.2. Definitions, Measures, and Analysis

Generally, cisgender refers to the alignment of one's internal sense of self and the social expectations associated with one's assigned sex at birth, whereas transgender or nonbinary is understood as having an internal sense of self that does not align with the social expectations associated with one's assigned sex at birth. We defined a patient in the sample as cisgender if their GI and ASAB aligned (e.g., male and male) or, when the ASAB was missing or unknown, if their GI and legal sex aligned (e.g., female and female). We designated a patient transgender if their GI and ASAB fields differed in a meaningful way (e.g., male and female; other and female) or, if the ASAB field was unknown, we also deferred to the legal sex (e.g., legally female with male gender identity). While some trans persons may be hidden in cases counted as cisgender—for example, where ASAB is missing and GI and legal sex align—we conversely felt it was important to count patients as trans whose ASAB may have been missing but GI and legal sex did not align.

The genesis of the two-step question was to improve the accuracy of the census of trans patients. In particular, this method attempts to count trans people who may simply identify as men or women and do not explicitly identify as transgender [29]. In this analysis, if a GI field indicates "Other" or "Nonbinary," we do not need to know their ASAB or legal sex to determine GI, as these two GIs indicate a gender that is explicitly outside the male/female gender binary. Therefore, if "Nonbinary" GI was selected with any other ASAB field or a null ASAB field, we counted the individual as nonbinary; similarly, "Other"

GI and any ASAB field counted as transgender, “Other.” “Choose not to disclose” was the singular gender identity response option that was not counted as cis or trans but was not excluded from the cohort like the null values (i.e., empty cells).

Patient characteristics related to GI, ASAB, and SO were assessed and compared between legally female and legally male patients. No statistical tests were applied to these distributions due to imbalance and small cell sizes. Comparing RUMC trans/nonbinary patients vs. cisgender patients, Chi-Square tests were used to assess the demographic and health-related distributions (i.e., age ranges, race, ethnicity, payor, smoking status). Smoking status, collected in the social history section where SO/GI data were collected prior to 2020, was the only health outcome variable for comparison within the available dataset.

Data quality checks were conducted via manual chart review in the following three domains: (1) ostensible GI errors: we manually validated charts with combinations of legal sex, ASAB, and GI that suggested a clear misclassification (e.g., a legal male, assigned male at birth, with GI listed as trans male), and (2) a randomly selected 10% of the final trans sample, and (3) a randomly selected 10% of the 41 charts with errors in SO fields – where GI data appeared. Only in the case of GI misclassification errors did we check all cases and reclassify them for this analysis based on manual review. Notes, medication orders, procedure codes, and keywords were used to shift a GI status from transfeminine to transmasculine or in the other direction. To determine that a chart was that of a cisgender patient, a chart was absent of transgender-related keywords, procedures, and medications, and we also were able to find keywords, medication, conditions, or chosen names that might have led a provider staff to recode a chart from cisgender to trans, and/or we were able to find official documentation in the chart that contradicted the classification (e.g., a patient classified as “legal F” whose chart has a picture of their driver’s license and sex indicating “M”). For the quality check of the trans sample we evaluated for obvious misclassification errors or for any additional confirmation of trans status, but if no confirming data were found, we did not consider it a misclassification.

To estimate the trans and nonbinary patient population prevalence at RUMC, patient data were excluded if the ASAB and GI fields were empty. The study was deemed exempt from review as human subjects research and was approved as quality improvement by the Rush Institutional Review Board.

3. Results

3.1. Sample Characteristics

The 2020 cohort of unique patients, ages 12 years and older, who had unplanned hospital encounters, consists of 49,314 individuals, and 24% ($n = 11,943$) of them had GI data fields completed while 76% ($n = 37,371$) remain empty (see Table 1; see Table A1 in Appendix A for more granular GI-related data). Thirty-six patients with GI data had “Choose not to disclose” as the GI response option and were set aside as neither cisgender nor transgender/nonbinary. Nine patients had legal sex, ASAB, and GI fields misaligned in a way that suggested they might be transgender but did not make logical sense; manual validation determined that three are trans and six are cisgender (see Table 2). In addition to the nonbinary (1) and other (2) patients, seven patients had transmasculine or transfeminine gender identities and unknown ASAB. In sum, nearly 1% ($n = 100$) of unique patients with GI data had data that reflected a transgender or nonbinary status.

Table 1. Gender identity data organized by assigned sex at birth (ASAB) for all Rush University Medical Center patients, 12 years and older, with unplanned hospital encounters in 2020 (N = 49,314).

Gender Identity	Assigned Sex at Birth				Total
	Female	Male	Nondisclosed	Unknown	
Man	28	3968		244	4240
Woman	7220	33	4	378	7635
Nonbinary	21	6		1	28
Other	1	1		2	4
Nondisclosed	9	7	11	9	36
Unknown	250	173	1	36,947	37,371
Total	7529	4188	16	37,581	49,314

Table 2. Legal sex, assigned sex at birth, and gender identity data ($n = 9$) that were illegible as cisgender or transgender status among the 2020 cohort of unique patients with unplanned admissions.

Error Labels		Corrected Labels			
Illegible Gender Identity		Cisgender		Transgender	
Legal Sex/ASAB/GI	$n = 9$	F/F/F	M/M/M	F/M/F	M/M/F
F/M/M	4	2	1	1	
M/M/TM	2				2
M/F/F	3	2	1		

For sexual orientation, 80% ($n = 39,862$) of data fields were unpopulated (see Table 3). No patients' SO data field contained the response option "Choose not to disclose." Although these are not response options in Epic's Demographic SO section, 13 patients had an SO data field indicating a "transgender" sexual orientation, and 28 patients had an SO data field indicating "neither exclusively male or female" as their sexual orientation. Excluding those 41 patients, 3.4% ($n = 322$) of all patients with populated SO data fields expressed a lesbian, gay, or bisexual sexual orientation. If we added the 28 nonbinary patients, who ostensibly have a non-heterosexual SO by virtue of being nonbinary, the prevalence of LGB patients would increase to 3.7%.

Table 3. Sexual orientation data for all Rush University Medical Center patients, 12 years and older, with unplanned hospital encounters in 2020 (N = 49,314).

Sexual Orientation	Legal F	Legal M	TOTAL
Gay	11	111	122
Heterosexual	5696	3392	9088
Bisexual	67	37	104
Lesbian	96	0	96
Transgender *	11	2	13
Neither man nor woman *	19	9	28
Choose not to disclose	0	0	0
Null/missing	22,087	17,775	39,862
Total by legal sex	27,987	21,326	49,313

* Does not represent a response option to sexual orientation in EHR; represents mismatching of gender ID data into sexual orientation fields.

The distributions of trans/nonbinary patients differed from cisgender ones across all demographic domains except for ethnicity (i.e., Latinx, non-Latinx, or unknown). Of note, 62% of trans/nonbinary patients were between ages 18–34, compared to 30% of cisgender ones; 40% of cisgender patients were age 55 years or older compared to 8% of trans/nonbinary patients (see Table 4). Smoking status also differed with trans/nonbinary patient data indicating 22% currently smoke, compared to 10% of cisgender patients.

Table 4. Demographic traits and smoking status of unique hospital patients (\geq age 12) in 2020, comparing patients with trans/NB GI data ($n = 100$) to patients with cisgender GI data ($n = 11,806$).

	Trans/Nonbinary		Cisgender		df, X ²	p-Value
	n = 100	%	n = 11,806	%		
Age						
12–17	6	6%	246	2%	4, 66.79	<0.001
18–34	62	62%	3535	30%		
35–44	13	13%	1753	15%		
45–54	11	11%	1596	14%		
55+	8	8%	4676	40%		
Assigned Sex at Birth						
Male	40	40%	3969	34%	2, 7.49	0.02
Female	50	50%	7219	61%		
ASAB unknown	10	10%	618	5%		
Race						
Black	24	24%	4349	37%	5, 11.52	0.04
White	54	54%	4866	41%		
Other	14	14%	2029	17%		
Asian/Native Hawaiian	5	5%	382	3%		
Native American	1	1%	45	0%		
Declined to answer/unknown	2	2%	135	1%		
Ethnicity						
Hispanic/Latinx	16	16%	2553	22%	2, 5.19	0.07
Non-Hispanic/non-Latinx	82	82%	9186	78%		
Declined/unknown	2	2%	67	1%		
Payor						
Managed care/Medicaid	49	49%	4777	40%	3, 23.31	<0.001
Blue Cross/Private	38	38%	3684	31%		
Medicare	5	5%	2648	22%		
None documented	8	8%	452	4%		
Other payors	0	0%	204	2%		
Self-pay	0	0%	41	0%		
Smoking Status						
Current	22	22%	1213	10%	3, 43.42	<0.001
Former	18	18%	2967	25%		
Never	52	52%	7446	63%		
Null/not documented	8	8%	180	2%		

3.2. Data Quality

Nine patients had combinations of legal sex, ASAB, and GI data that were illegible in terms of a clear cisgender or transgender determination (see Table 2). In sum, after manual chart reviews, six patients were deemed cisgender and three transgender. Among one grouping, four patients were erroneously labeled legal female, assigned male at birth, and male GI (i.e., F/M/M). Three of these patients were cisgender, two women and one man. The man's legal sex was miscoded as "F" and the women's ASAB and GI were miscoded "M." One of the cisgender women patients had her husband's first name listed as her "preferred name," and she takes a medication associated with transgender women, which may have led RUMC staff to misinterpret the patient's chart as transgender and recode the patient as F/M/M. The fourth patient with data labeled F/M/M was determined to be transfeminine based on evidence of a legal feminine name, alternating misgendering and appropriate gendering by providers in the notes, along with transgender keywords in the notes, and gender-affirming hormones in the medication list.

Ten patients comprised the random 10% data quality check among the subsample of trans patients. Of these, five had additional information in the chart supporting trans

status. For example, a patient deemed transgender male according to GI and ASAB data also had testosterone in his prescription medication list. Of the five patients with information in the chart supporting transgender status, three were transfeminine and two were transmasculine. Four patients had no supporting or contradicting information regarding transgender status. The tenth patient's gender identity data had been changed between our initial data analysis and the subsequent data quality check from transgender to cisgender female. The procedural history for this patient is consistent with cisgender status.

Although sexual orientation data collection is not the primary focus here, we conducted a quality check on 10% of the 41 patients with gender identity-related data listed as their sexual orientation. As Figure 1 demonstrates, gender identities are not response options for sexual orientation. The chart review of one "transgender" case (i.e., patient Doe) and three "neither exclusively male or female," revealed that all four charts had GI data in the EHR's social history section, which—prior to 2020 when the SO/GI demographics SmartForm was deployed—was the primary place for SO/GI data collection. It appears that these data have been mismapped to sexual orientation, sometimes even overriding existing SO data that appears in either social history or the demographic field for sexual orientation. Three of four of these patient charts had lesbian, gay, and heterosexual sexual orientations in the SO field, and the fourth was empty. Instead, for these patients ($n = 41$), SmartPhrases related to gender identity like the Epic definition provided for nonbinary, "neither exclusively man or woman" (see Figure 1) were mapped to sexual orientation in our report.

4. Discussion

To our knowledge, this analysis is the first to benchmark implementation penetration and quality of gender identity-related data collection in U.S. academic health centers since the Meaningful Use incentivization for this data collection went into effect in 2014. Using structured data for gender identity and assigned sex at birth, 100 patients, or 0.84%, of the 2020 unplanned hospital patient cohort with GI data were identified as TGNB, in comparison to the estimated 0.5% of adults in Chicago in 2018, 0.51% of adults in Illinois in 2014, and 0.5–0.6% of adults in the United States in 2014 [30,31]. Compared to the cisgender hospital patients with gender ID data in 2020, the trans patients were disproportionately young, white, Medicaid insured, and had a higher prevalence of smoking. The disparities between the cis and trans samples indicate the critical importance of both SO/GI data collection and learning health systems science vis-à-vis equitable access to care, utilization outcomes, and health outcomes [32].

Our real-world data were too sparse and of limited quality to draw population-based conclusions. Although Rush is unique and a leader for its active and routine collection of SO/GI data [17], 76% of the EHR GI data fields remain unpopulated and highlight the need to address SO/GI data collection implementation appropriateness, penetration, and effectiveness more widely. Rush should continue to lead other systems and improve on the rate of SO/GI data collection and penetration with continued training efforts across hospital departments and at all levels within those departments. Among the populated fields of ASAB and GI as well as SO, the data quality requires improvement for more meaningful learning health systems research and clinical decision making. Widespread training regarding gender-affirming care and inclusive language and the development of SO/GI data quality standards will move U.S. healthcare closer to the goals of the Meaningful Use mandate. Patient Doe's SO/GI data exemplify two key issues associated with our data quality checks: (1) the systemic error whereby some patients have GI data in their SO data fields indicates low SO/GI literacy among provider staff, in this case, who manage EHR data, and (2) the low literacy, alongside TGNB stigma, discrimination, and violence in U.S. healthcare settings may contribute to patient SO/GI disclosure choices.

Our QI study is not without limitations. In 2020, COVID-19 limited the number of non-COVID-19 related emergencies and unplanned admissions [33] biasing the sample compared to prior years. Selection bias also likely impacted our cohort whereby those

patients with missing gender identity data are likely different from those patients who have GI data in their charts. Both provider and patient levels of literacy with concepts of gender identity are likely associated with whether or not a GI-related field is populated. Similarly, both the provider and patient's comfort levels with discussion of gender identity and its relationship to health are likely associated with the presence or absence of GI-related data in the EHR. With unplanned hospital admissions, the severity and acuity of illness also matter and may influence provider perception of the appropriateness of SO/GI data collection in an acute care setting. Providers may have less time or concern to collect GI-related data from very sick patients.

5. Conclusions

Electronic health record data are not purely infrastructural [34]. Encounter data reflect moments of intense vulnerability for patients, particularly TGNB patients. These vulnerabilities are shaped not only by levels of patient health literacy and provider SO/GI literacy but also by the lack of durable legal protections in the U.S. that govern TGNB access to care, insurance coverage, and data privacy. Protections issued by executive order and rules, such as Section 1557 of the Affordable Care Act, that may change with each administration, limit acceptability and effectiveness of SO/GI data collection and disclosure in the United States [35,36]. Without more durable support and protections, SO/GI data collection will have a threshold that is fractional and fraught [27].

Rule-based algorithms and computable phenotypes that draw on a broad spectrum of structured data including diagnostic codes, procedure codes, medication orders, as well as a range of transgender and nonbinary-related keywords labeled in clinical notes have been generated to identify more comprehensive TGNB patient cohorts and associated health outcomes [23,37–39]. Machine learning and natural language processing (NLP) techniques may also operationalize critical information in clinical notes, rectifying current TGNB population health research gaps and relieving pressure on both patients and providers to disclose or collect SO/GI demographic data [40]. As with structured data collection, innovations to improve equity in gender-affirming care, research, and health outcomes will likely be more effective with a learning health systems science approach that includes ongoing provider training, patient education, and community and stakeholder engagement.

Author Contributions: Conceptualization, C.A.K. Study design, H.M.T. Methodology, H.M.T., P.P., C.A.K. Data curation, P.P., H.M.T. Software, P.P., H.M.T. Validation, H.M.T., K.F. Formal analysis, H.M.T., P.P. Investigation, H.M.T., K.F., P.P. Resources, N.S.K. Writing—original draft, H.M.T. Writing—review and editing, C.A.K., N.S.K., K.F., H.M.T. Visualization, H.M.T., K.F., C.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: Hale Thompson's contribution to this project is supported by the Agency for Healthcare Research & Quality grant number K12-HS-026385 and the National Institute for Drug Abuse grant number 5R25DA035692-08. The content is solely the responsibility of the authors and does not represent the official views of the Agency for Healthcare Research and Quality nor those of the National Institute for Drug Abuse.

Institutional Review Board Statement: Ethical review and approval were waived for this study, as the Rush University Medical Center Research Regulatory Operations Office determined it to be a quality improvement study rather than research or human subjects research (Record ID 2930).

Informed Consent Statement: As a quality improvement study, informed consent was not required and would not be practical given the scope of the analysis.

Data Availability Statement: Our dataset is derived from patient electronic health records, which include protected health information, and we cannot make it publicly available due to regulatory and legal restrictions imposed by Rush University Medical Center. Patient medical data is highly sensitive and, with quasi-identifiers such as race, ethnicity, and age, medical record data is reidentifiable when linked to other publicly available datasets. Should researchers who meet the criteria for access to this confidential data want to use our de-identified dataset to replicate the gender identity data collection implementation adoption and quality assessment, our Chief of Research Informatics at Rush (Casey

Frankenberger) will serve as the point of contact outside our research team to inquire further about establishing a data use agreement and access to the patient data (cfranken@rush.edu).

Acknowledgments: The authors would like to acknowledge Ekta Kishen, MS for her contributions to this analysis and all the transgender and nonbinary patients who use a range of strategies to seek and receive care.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Assigned sex at birth, gender identity, and legal sex EHR data for all Rush University Medical Center patients, 12 years and older, with unplanned hospital encounters in 2020 ($N = 49,314$).

	Legal F	Legal M	TOTAL
Assigned Sex			
Female	7511	18	7529
Male	23	4165	4188
Choose not to disclose	14	8	22
Not recorded on BC	2	0	2
Unknown/uncertain	4	4	8
Null/missing	20,433	17,131	37,564
Gender Identity			
Trans woman (AMAB)	8	8	16
Trans man (AFAB)	12	3	15
Cis woman (AFAB)	7220		7220
Cis man (AMAB)		3968	3968
Non-binary (AFAB)	19	2	21
Non-binary (AMAB)	0	6	6
Non-binary (unknown ASAB)	0	1	1
Woman (AMAB)	9	8	17
Man (AFAB)	4	9	13
Trans woman (unknown ASAB)		2	2
Trans man (unknown ASAB)		2	2
Woman (unknown ASAB)*	376		376
Man (unknown ASAB)*	3	239	242
Other (AFAB)	1		1
Other (AMAB)	1		1
Other (unknown ASAB)	2		2
Choose not to disclose	26	10	36
Null/missing	20,307	17,067	37,374
Legal Sex			
Female	27,987		27,987
Male		21,326	21,326
Unknown			1
TOTAL			49,314

References

- Collin, L.; Reisner, S.L.; Tangpricha, V.; Goodman, M. Prevalence of Transgender Depends on the “Case” Definition: A Systematic Review. *J. Sex. Med.* **2016**, *13*, 613–626. [CrossRef]
- Thompson, H.; King, L. Who Counts as “Transgender”? Epidemiological Methods and a Critical Intervention. *TSQ Transgend. Stud. Q.* **2015**, *2*, 148–159. [CrossRef]
- Dutta, A. Between Aid Conditionality and Identity Politics—The MSM-Transgender Divide and Normative Cartographies of Gender vs. Sexuality: Aniruddha Dutta. Kafila.org. 2012. Available online: <http://kafila.org/2012/04/05/between-aid-conditionality-and-identity-politics-the-msm-transgender-divide-and-normative-cartographies-of-gender-vs-sexuality-aniruddha-dutta/#more-12257> (accessed on 10 June 2021).
- Hammack, P.L.; Hughes, S.D.; Atwood, J.M.; Cohen, E.M.; Clark, R.C. Gender and Sexual Identity in Adolescence: A Mixed-Methods Study of Labeling in Diverse Community Settings. *J. Adolesc. Res.* **2021**. [CrossRef]
- UNESCO. *Behind the Numbers: Ending School Violence and Bullying*; UNESCO Publishing: Paris, France, 2019.

6. Dutta, A. Legible identities and legitimate citizens. *Int. Fem. J. Politics* **2013**, *15*, 494–514. [CrossRef]
7. Dutta, A.; Roy, R. Decolonizing transgender in India: Some reflections. *TSQ Transgender Stud. Q.* **2014**, *1*, 320–337. [CrossRef]
8. Hanssmann, C. Epidemiological rage: Population, biography, and state responsibility in trans- health activism. *Soc. Sci. Med.* **2020**, *247*, 112808. [CrossRef]
9. Ashley, F. “Trans” Is My Gender Modality: A Modest Terminological Proposal. Available online: https://www.florenceashley.com/uploads/1/2/4/4/124439164/florence_ashley_trans_is_my_gender_modality.pdf (accessed on 10 June 2021).
10. World Health Organization. International Classification of Diseases (ICD); Published Online First: 2006. Available online: <https://icd.who.int/browse10/2019/en> (accessed on 10 June 2021).
11. World Health Organization. International classification of diseases—Ninth revision (ICD-9). *Wkly. Epidemiol. Rec. Relev. Épidémiol. Hebd.* **1988**, *63*, 343–344.
12. Khoury, B.; Kogan, C.; Daouk, S. International Classification of Diseases 11th Edition (ICD-11). *Encycl. Personal. Individ. Differ.* **2020**, 2350–2355. [CrossRef]
13. World Professional Association for Transgender Health. Standards of Care for the Health of Transsexual, Transgender, and Gender-Conforming People [7th Version]. 2012. Available online: <https://www.wpath.org/publications/soc> (accessed on 10 June 2021).
14. Jones, T. Researching and Working for Transgender Youth: Contexts, Problems and Solutions. *Soc. Sci.* **2016**, *5*, 43. [CrossRef]
15. UNESCO. *From Insult to Inclusion: Asia-Pacific Report on School Bullying, Violence and Discrimination on the Basis of Sexual Orientation and Gender Identity*; UNESCO: Paris, France, 2015.
16. Smith, E.; Jones, T.; Ward, R.; Dixon, J.; Mitchell, A.; Hillier, L. *From Blues to Rainbows: The Mental Health and Well-Being of Gender Diverse and Transgender Young People in Australia*; Australian Research Centre in Sex, Health and Society (ARCSHS): Bundoora, Australia, 2014.
17. The Human Rights Campaign. Healthcare Equality Index. Washington D.C.: Human Rights Campaign 2019. Available online: http://www.hrc.org/hei/leaders-in-lgbt-healthcare-equality#.VOflw_nF-Sr (accessed on 10 June 2021).
18. Grasso, C.; Goldhammer, H.; Funk, D.; King, D.; Reisner, S.L.; Mayer, K.H.; Keuroghlian, A.S. Required Sexual Orientation and Gender Identity Reporting by US Health Centers: First-Year Data. *Am. J. Public Health* **2019**, *109*, 1111–1118. [CrossRef]
19. Plemons, E. A Capable Surgeon and a Willing Electrologist: Challenges to the Expansion of Transgender Surgical Care in the United States. *Med. Anthropol. Q.* **2019**, *33*, 282–301. [CrossRef]
20. Thompson, H.M. Stakeholder Experiences with Gender Identity Data Capture in Electronic Health Records: Implementation Effectiveness and a Visibility Paradox. *Health Educ. Behav.* **2021**, *48*, 93–101. [CrossRef] [PubMed]
21. Thompson, H.M. Patient Perspectives on Gender Identity Data Collection in Electronic Health Records: An Analysis of Disclosure, Privacy, and Access to Care. *Transgend. Health* **2016**, *1*, 205–215. [CrossRef] [PubMed]
22. Cruz, T.M. The making of a population: Challenges, implications, and consequences of the quantification of social difference. *Soc. Sci. Med.* **2017**, *174*, 79–85. [CrossRef] [PubMed]
23. Kronk, C.; Everhart, A.R.; Ashley, F.; Thompson, H.M.; Schall, T.E.; Goetz, T.G.; Hiatt, L.; Derrick, Z.; Queen, R.; Ram, A. Transgender data collection in the electronic health record (EHR): Current concepts and issues. *JAMIA* **2021**. [CrossRef]
24. Callahan, E.J.; Sitkin, N.; Ton, H.; Eidson-Ton, W.S.; Weckstein, J.; Latimore, D. Introducing Sexual Orientation and Gender Identity Into the Electronic Health Record: One Academic Health Center’s Experience. *Acad. Med.* **2015**, *90*, 154–160. [CrossRef]
25. Lewis, C.C.; Fischer, S.M.; Weiner, B.J.; Stanick, C.; Kim, M.; Martinez, R.G. Outcomes for implementation science: An enhanced systematic review of instruments using evidence-based rating criteria. *Implement. Sci.* **2015**, *10*, 155. [CrossRef]
26. Razzaghi, H.; Greenberg, J.; Bailey, L.C. Developing a systematic approach to assessing data quality in secondary use of clinical data based on intended use. *Learn. Health Syst.* **2021**. [CrossRef]
27. Weng, C. Clinical data quality: A data life cycle perspective. *Biostat. Epidemiol.* **2020**, *4*, 6–14. [CrossRef]
28. Bourgois, P.; Holmes, S.M.; Sue, K.; Quesada, J. Structural Vulnerability: Operationalizing the Concept to Address Health Disparities in Clinical Care. *Acad. Med.* **2017**, *92*, 299–307. [CrossRef]
29. Tate, C.C.; Ledbetter, J.N.; Youssef, C.P. A Two-Question Method for Assessing Gender Categories in the Social and Medical Sciences. *J. Sex. Res.* **2013**, *50*, 767–776. [CrossRef]
30. Weaver, K.N.; Laflamme, E.; Prachard, N. *Lesbian, Gay, Bisexual & Transgender Health. Healthy Chicago Databook*; Chicago Department of Public Health: Chicago, IL, USA, 2018.
31. Flores, A.R.; Herman, J.L.; Gates, G.J.; Brown, T.N.T. *How Many Adults Identify as Transgender in the United States?* The William Institute: Los Angeles, CA, USA, 2016.
32. Greene, S.M.; Reid, R.J.; Larson, E.B. Implementing the learning health system: From concept to action. *Ann. Intern. Med.* **2012**, *157*, 207–310. [CrossRef] [PubMed]
33. Birkmeyer, J.D.; Barnato, A.; Birkmeyer, N.; Bessler, R.; Skinner, J. The Impact of the COVID-19 Pandemic On Hospital Admissions In The United States. *Health Aff.* **2020**, *39*, 2010–2017. [CrossRef] [PubMed]
34. Crawford, K. *The Atlas of AI*; Yale University Press: London, UK, 2021.
35. Jones, T. Trump, trans students and transnational progress. *Sex Educ.* **2018**, *18*, 479–494. [CrossRef]
36. Section 1557 of the Patient Protection and Affordable Care Act | HHS.gov. 2010. Available online: <http://www.hhs.gov/civil-rights/for-individuals/section-1557/index.html> (accessed on 18 June 2021).

37. Guo, Y.; He, X.; Lyu, T.; Zhang, H.; Wu, Y.; Yang, X.; Chen, Z.; Markham, M.J.; Modave, F.; Xie, M.; et al. Developing and validating a computable phenotype for the identification of transgender and gender nonconforming individuals and subgroups. *bioRxiv* **2020**. [[CrossRef](#)]
38. Ehrenfeld, J.M.; Gottlieb, K.G.; Beach, L.; Monahan, S.E.; Fabbri, D. Development of a Natural Language Processing Algorithm to Identify and Evaluate Transgender Patients in Electronic Health Record System. *Ethn. Dis.* **2019**, *29*, 441–450. [[CrossRef](#)]
39. Kronk, C.; Tran, G.Q.; Wu, D.T.Y. Creating a Queer Ontology: The Gender, Sex, and Sexual Orientation (GSSO) Ontology. *Stud. Health Technol. Inform.* **2019**, *264*, 208–212.
40. Lynch, K.E.; Viernes, B.; Schliep, K.C.; Gatsby, E.; Alba, P.R.; DuVall, S.L.; Blossnich, J.R. Variation in Sexual Orientation Documentation in a National Electronic Health Record System. *LGBT Health* **2021**, *8*, 201–208. [[CrossRef](#)]