Case Reports (JAMIA)

ABSTRACT

Since 2017, the TrueNTH Global Registry (TNGR) has aimed to drive improvement in patient outcomes for individuals with localized prostate cancer by collating data from healthcare institutions across thirteen countries. As TNGR matures, a systematic evaluation of existing processes and documents is necessary to evaluate whether the registry is operating as intended.

The main supporting documents: protocol and data dictionary, were comprehensively reviewed in a series of meetings over a 10-month period by an international working group. In parallel, individual consultations with local institutions regarding a benchmarking quality-of-care report were conducted. Four consensus areas for improvement emerged: updating operational definitions, appraisal of the recruitment process, refinement of data elements, and improvement of data quality and reporting. Recommendations presented were drawn from our collective experience and accumulated knowledge in operating an international registry. These can be readily generalized to other health-related reporting programs beyond clinical registries.

Case Reports (JAMIA)

INTRODUCTION

Clinical Quality Registries (CQRs) have the ability to facilitate improvement of clinical care and patient outcomes through performance benchmarking[1 2]. Rigorous documentation and evaluation of processes increase consistency and strengthen a CQR's capacity to meet these objectives, while inadequacies in key operational documents and processes hinder effective and efficient operations. Unless addressed, adverse consequences on healthcare providers, patients, funders, and other stakeholders may ensue[3-5]. A CQR should also remain cognizant of contemporary clinical practices. Practice changes that develop during a registry program may necessitate an expansion of the dataset or ceasing the collection of certain data elements.

CASE DESCRIPTION

The TrueNTH Global Registry (TNGR) was initiated in 2017 to monitor and optimize quality of prostate cancer care by systematically collecting data from prostate cancer registries/research projects worldwide[6]. Comparative benchmark reports assessing indicators of quality care are produced and released to participating sites biannually, and are an integral part of this international registry[7]. The aim of this report is to share examples and learnings from a collaboration that was formed to review key processes supporting TNGR.

MATERIALS AND METHODS

Two parallel cooperative approaches were established. First, an advisory working group (WG) was formed to undertake a holistic review of the main documentation tools supporting TNGR: the protocol and data dictionary, and a major registry output: the benchmarking quality-of-care report (Figure 1). Second, individual consultations with local data centers ("sites") focused on their tailored quality-of-care report.

[Figure 1 is about here]

Case Reports (JAMIA)

At the direction of the TNGR Executive Committee (EC), expressions of interest for the WG were requested from all stakeholders. An overarching goal in constituting the WG was to have diversity of voices, adequate geographic representation, and to build notions of ownership in the TNGR. Table 1 lists the full membership of the core team and the advisory WG, including their country of practice.

Table 1: Background of the core project team and advisory working group

	Urology	Radiation Oncology	Psychology	Epidemiology	Academic	Patient Representative	Registry Manager	TOTAL
Australia	-	1	-	-	1*	-	1* + 1	4
Germany	-	-	-	-	1*	1	1	3
Hong Kong	1	-	-	-	-	-	1	2
Spain	-	-	-	1		-	1	2
UK	-	-	1	-	-	-	-	1
US	-	-	-	1*	-	-	1* + 1	3
TOTAL	1	1	1	2	2	1	7	15

* Core project team

A series of monthly one-hour videoconference consultation sessions were held over a 10-month period. Central issues considered at meetings related to the protocol were the continuing relevance of each section, clarity of language, and new aspects of registry operations that needed incorporation. The data dictionary review meetings were oriented to the interpretability of data elements, feasibility of collection, and whether they're still relevant to contemporary practice. The WG was also encouraged to identify proposed new data items supported by literature reviews.

[Figure 2 is about here]

Regarding appraising of the quality-of-care report, the TNGR Data Coordination Centre (DCC) ran draft reports for participating sites, examined each and noted any anomalies. Principal investigators and data managers were asked to provide responses to the issues raised and/or necessitate action

within four weeks to allow the opportunity to remedy these prior to final report dissemination. Recommendations from this activity informed future refinements.

RESULTS

The WG meetings and subsequent iterative consultations uncovered four main areas for improvement. Each is illustrated with examples and accompanied by background context and learning points.

1 | UPDATING OPERATIONAL DEFINITIONS

Example 1(a) | Defining de-identification techniques according to evolving legal standards

Data from sites contributing to TNGR are transferred biannually to a central repository. Anonymity of registry participants is preserved with unique identifiers provided by each site[8]. This key identifier is held in their local registry and can be linked to the full patient data for error correction and follow-up, without fully identified information being released to TNGR.

Consultation was held with the DCC legal counsel to establish an internationally acceptable standard definition to describe the mechanism of re-identification keys used in TNGR. Pseudonymization of the participant identifier was determined to be the most precise term based on legal definitions[9]. Briefly, this technique requires that identifiable personal data are kept separate and replaced by an indirect identifier or a pseudonym so the data can no longer be attributed to a specific registry participant[10].

Learning points: As the TNGR protocol was prepared before newer European guidelines came into effect, synchronization of terminology with up-to-date regulatory guidelines is necessary to provide assurance that the CQR operates in alignment with worldwide technical and operating standards.

Example 1(b) | Clarifying the definition of disease stage for case ascertainment

The term "locally advanced" was in the eligibility criteria for TNGR, but its interpretation among sites was found to vary. The phrase was adopted from older guidelines[11] and embedded in TNGR recruitment processes from the outset of operation. In line with revised guidelines, WG consensus was reached to apply the term "regional" to designate prostate cancer with regional lymph node spread (N1), but with no distant metastases (M0)[12].

Learning points: Retention of imprecise terms and variable interpretations by registry personnel affects recruitment and hinders accurate cross-site comparisons.

2 | APPRAISAL OF THE RECRUITMENT PROCESS

Example 2(a) | Revising the recruitment process at referral centers

Disproportionately high volumes of missing treatment information were observed at referral centers. Numerous "consult only" patients were enrolled to TNGR, although they had not been diagnosed nor received treatment/surveillance at those centers. Consequently, there were high number of patients "lost to follow-up".

Lack of information integral to measuring clinical quality impairs TNGR's capacity to monitor practice and conduct systematic analysis comparing site performance. The EC further concurred that participating sites should only recruit individuals who were diagnosed and/or received treatment for prostate cancer at their center. Including patients who only presented for consultation may result in data unrepresentative of the center. This issue has stimulated the centers to reappraise their recruitment process.

Learning points: Recruiting unbiased patient samples with high response rates is the goal of a CQR. The inclusion of individuals by centers contributing to a CQR should be representative of national,

subnational or institution cases. Nevertheless, the decision to exclude specific individuals from a CQR must be carefully weighed against the risk of selection bias.

3 | REFINEMENT OF DATA ELEMENTS

Example 3(a) | Refining difficult to collect data elements

The collection of ethnicity and socioeconomic status is problematic in an international setting. The ways of recording these are heterogenous which limits their use in analysis even though they are valuable covariates. Recommendations from published studies were generally uninformative, as they were confined to a state or nationwide population[13-15].

The WG decided to retain ethnicity as an optional, free-text field, deferring to local or country standards for its definition and categorization. This may permit intra-jurisdictional analysis (if those sites are coherent) but precludes registry-wide harmonization. Socioeconomic status collected as a free-text field was identified as not fit for purpose. A surrogate approach was found to collect instead *the highest level of education*, defined according to a modified version of the International Standard Classification of Education[16].

Learning points: The inclusion of ethnicity and socioeconomic status in an international registry remains challenging and a pragmatic approach to its collection and use is necessary.

Case study 3(b) | Adding a data element

There is an emerging data supporting the use of multiparametric Magnetic Resonance Imaging and the Prostate Imaging-Reporting and Data System (PI-RADS) to guide management of localized prostate cancer[17-19]. Inclusion of this item was proposed by a WG member during data dictionary review and its collection was deemed to be feasible by a site survey. The addition has a minimal impact on TNGR's current quality assurance initiatives as risk categorization in many jurisdictions

does not yet include PI-RADS, however, the registry is now well placed to evaluate prospective treatment paradigms.

Learning points: The inclusion of PI-RADS bridges a gap between TNGR and real-world clinical practice. A registry should continually assess whether their data elements are evidence-based, have clinical utility in multiple jurisdictions, and high feasibility of accurate collection. Proposals for new items can regularly be sought but a rigorous process for their inclusion must be codified.

Example 3(c) | Accommodating a new option to an existing data element

The occurrence of non-detectable carcinoma in a prostatectomy specimen (pT0) after positive biopsy is unusual, with an incidence of 0.4%[20]. pT0 did not fit in the pathological tumor stage (pT) categories defined by AJCC 8th edition, which ranges from pT2 to pT4[19], however, considering this stage occurs in TNGR, a decision was made to add "pT0" as a new category.

Learning points: A CQR should address valid incongruities detected after cases begin accruing and incorporate practical solutions to collect them.

4 | IMPROVEMENT OF DATA QUALITY AND REPORTING PROCESSES

Example 4(a) | Handling high missingness of data elements used in quality indicators

Collection of data elements such as radiology imaging was optional in TNGR, because they were not part of the ICHOM Standard Set[21]. However, they are recognized as clinically relevant variables[6 8] and form part of some quality indicator calculations[22]. The magnitude of missing data impacted the number of cases that could be reasonably evaluated within a site.

Two recommendations were proposed and accepted. First, prescribe a stringent criterion where inclusion of a site in a particular benchmarked indicator is warranted only if "completeness" reached

a sufficient threshold. Second, compel a site to collect data elements which were used in quality indicators' computation. An 80% completeness cut-off was proposed considering that: 1) it is the target level of completeness for collection of non-mandatory data elements set out in the TNGR protocol; 2) most centers (greater than 60%) can still be evaluated in the benchmark indicator.

Learning points: Excessive amounts of missing data confer a bias, misrepresent a site's performance and limit the validity of a benchmarked quality indicator. Setting a reasonable completeness threshold for inclusion of sites in an indicator will facilitate greater accuracy and, in turn, drive improvements. Changing optional data elements to mandatory theoretically incentivizes data completeness. However, the decision whether to upgrade data elements from "optional" to "mandatory" can have a critical impact on allocation of resources at a site. This decision establishes a precedent for future participating institutions.

Example 4(b) | Consolidating registry entries and avoiding duplicates

Two identifiers were used to uniquely identify a registry participant in TNGR. *LDCCode* which distinguished sites within countries, and a unique identifier associated with individuals within a local center (*LDCPatientID*)[8]. During the review process, sites were reporting more participants than was expected due to individuals being entered more than once into the TNGR but with unique identifiers.

To minimize future duplicate entries being submitted to the central registry, a new process has been employed by the DCC. Sites are asked to verify individuals that were uploaded with the same set of numeric and date baseline variables against records in their local registry where participants' personal information was held. Confirmed duplicate entries are subsequently removed from TNGR.

Case Reports (JAMIA)

<u>Learning points</u>: When limited personal information is available, baseline numeric or date variables such as date of diagnosis, blood test result, etc. can be used as an initial flag for potential duplicate entries.

DISCUSSION

Standardization of operational definitions is particularly relevant to a CQR. Literacy in informatics concepts such as de-identification methodology among health/research community has been a documented concern[23 24], suggesting the need for education on this subject. Clinical terminologies selected to describe registry populations should be in line with current practice to promote harmonization.

There is a trade-off that registries face when considering inclusion of additional data elements. The benefits of having data elements which are of interest to our stakeholders impose a greater demand in time and financial resources to support collection[2]. This concern must be weighed against the risk of not collecting data that informs and improves patient outcomes. Continuing ineffective registry practices wastes resources which has negative downstream consequences.

Limited guidance exists in relation to the acceptable cut-off of proportion of missingness that warrant inclusion for reporting, ranging from 70%[5] to 95%[25]. Others indicated that missingness greater than 10% may bias statistical analysis[26]. A collective effort will be carried forward to assess the impact of our updates in participant management and quality measures calculation; and to enhance data quality which form the basis of the indicators.

While the examples presented in this report are not exhaustive, the specific recommendations were drawn from our collective experience and accumulated knowledge in running an international CQR. With more robust frameworks are in place, we can envisage wider collaborations moving forward.

CONCLUSIONS

We identified four major areas for improvement following a comprehensive and collaborative review of key documents supporting an international clinical registry. Such information and lessons learned are translatable to other programs that collect, manage, analyze, and report health data.

ETHICS APPROVAL

The conduct of this internal review process is a component of the TrueNTH Global Registry Prostate Cancer Outcomes, approved by the Alfred Health Human Research Ethics Committee (HREC/16/Alfred/98) Protocol version 2.0 (7 Sept 2017). Participation in the working group was voluntary.

REFERENCES

1. Wilcox N, McNeil JJ. Clinical quality registries have the potential to drive improvements in the appropriateness of care. Med J Aust. 2016 Nov 21;205(10):S27-S29. doi: 10.5694/mja15.00921. PMID: 27852198.
2. Stey AM, Russell MM, Ko CY, Sacks GD, Dawes AJ, Gibbons MM. Clinical registries and quality measurement in surgery: a systematic review. Surgery 2015;**157**(2):381-95 doi: 10.1016/j.surg.2014.08.097[published Online First: Epub Date]|.
3. Alexander M, Evans SM, Wolfe R, Ball DL, Burbury K. Risks of using medical record and administrative data for prognostic models. Med J Aust. 2017 Aug 7;207(3):126. doi: 10.5694/mja16.00919. PMID: 28764629.
4. D'Amore JD, McCrary LK, Denson J, et al. Clinical data sharing improves quality measurement and patient safety. J Am Med Inform Assoc 2021;**28**(7):1534-42 doi: 10.1093/jamia/ocab039[published Online First: Epub Date]|.
5. Naik S, Voong S, Bamford M, Smith K, Joyce A, Grinspun D. Assessment of the Nursing Quality Indicators for Reporting and Evaluation (NQuIRE) database using a data quality index. Journal of the American Medical Informatics Association 2020;**27**(5):776-82
6. Evans SM, Millar JL, Moore CM, et al. Cohort profile: the TrueNTH Global Registry - an international registry to monitor and improve localised prostate cancer health outcomes. BMJ Open 2017;**7**(11):e017006 doi: 10.1136/bmjopen-2017-017006[published Online First: Epub Date]|.
7. Sampurno F, Cally J, Opie JL, et al. Establishing a global quality of care benchmark report. Health Informatics J 2021;**27**(2):14604582211015704 doi: 10.1177/14604582211015704[published Online First: Epub Date]|.
8. Sampurno F, Kannan A, Lucas M, et al. Development of Technologic Solutions to Address Complex Local Requirements of an International Prostate Cancer Clinical Quality Registry. JCO Clin Cancer Inform 2019;**3**:1-11 doi: 10.1200/CCI.18.00114[published Online First: Epub Date]|.
9. Article 4(5) of the GDPR. https://gdpr-info.eu/art-4-gdpr/. Accessed September 2021.
10. Pung J, Rienhoff O. Key components and IT assistance of participant management in clinical research: a scoping review. JAMIA Open 2020;**3**(3):449-58 doi: 10.1093/jamiaopen/ooaa041[published Online First: Epub Date]|.
11. NCCN guidelines: prostate cancer Version 3, 2016. https://www2.tri-kobe.org/nccn/guideline/archive/urological2016-2017/english/prostate.pdf. Accessed September 2021.
12. NCCN guidelines: prostate cancer Version 4, 2019. https://www2.tri-kobe.org/nccn/guideline/urological/english/prostate.pdf. Accessed September 2021.
13. Ramsey SD, Zeliadt SB, Hall IJ, Ekwueme DU, Penson DF. On the importance of race, socioeconomic status and comorbidity when evaluating quality of life in men with prostate cancer. J Urol. 2007 Jun;177(6):1992-9. doi: 10.1016/j.juro.2007.01.138. PMID: 17509278.

Case Reports (JAMIA)

14. Morning, A. Ethnic Classification in Global Perspective: A Cross-National Survey of the 2000 Census Round. Popul Res Policy Rev 27, 239–272 (2008). https://doi.org/10.1007/s11113-007-9062-5.
15. NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database. https://www.cdc.gov/cancer/uscs/public-use/pdf/npcr-seer-public-use-database-data-dictionary-2001-2014.pdf. Accessed September 2021.
16. United Nations Educational, Scientific and Cultural Organization International standard classification of education: ISCED 2011 www.uis.unesco.org/Education/Documents/isced-2011-en.pdf. Accessed September 2021.
17. Mottet N, van den Bergh RC, Briers E, et al. . 2019 EAU - EANM - ESTRO - ESUR - SIOG Guidelines on Prostate Cancer, Vol 53. Arnhem, the Netherlands: EAU Guidelines Office, 2019.
18. National Institute for Health and Care Excellence (NICE). Prostate cancer: diagnosis and management (update). 2019. https://www.nice.org.uk/guidance/indevelopment/gid-ng10057. Accessed September 2021.
19. NCCN Clinical Practice Guidelines in Oncology V1 .2019. Prostate cancer early detection recommendations. J Natl Compr Canc Netw 2019.
20. Bream MJ, Dahmoush L, Brown JA. pT0 Prostate Cancer: Predictive Clinicopathologic Features in an American Population. Curr Urol 2013;**7**(1):14-8 doi: 10.1159/000343546[published Online First: Epub Date]|.
21. ICHOM Localised Prostate Cancer Version 2.0.5. Revised: April 7th, 2017. https://ichom.org/files/medical-conditions/localized-prostate-cancer/localized-prostate-cancer-reference-guide.pdf. Accessed September 2021.
22. Sampurno F, Zheng J, Di Stefano L, et al. Quality Indicators for Global Benchmarking of Localized Prostate Cancer Management. J Urol 2018;**200**(2):319-26 doi: 10.1016/j.juro.2018.02.071[published Online First: Epub Date]|.
23. Chevrier R,Foufi V,Gaudet-Blavignac C,Robert A,Lovis C. Use and Understanding of Anonymization and De-Identification in the Biomedical Literature: Scoping Review. J Med Internet Res 2019;21(5):e13484. doi: 10.2196/13484
24. Schraven SP, Mlynski R. Evaluation of Multicenter Registry Data. Laryngorhinootologie 2019;**98**(S 01):S173-S96 doi: 10.1055/a-0806-0971[published Online First: Epub Date]|.
25. National Program of Cancer Registries (NPCR) and Surveillance, Epidemiology & End Results (SEER) Incidence: USCS 2001 to 2014 Public Use Database Data Standards and Data Dictionary. November 2016 Submission, Diagnosis Years 2001 through 2014. (cdc.gov). https://www.cdc.gov/cancer/uscs/public-use/pdf/npcr-seer-public-use-database-data-dictionary-2001-2014.pdf. Accessed September 2021.
26. Bennett DA. How can I deal with missing data in my study? Aust N Z J Public Health. 2001 Oct;25(5):464-9. PMID: 11688629.