Enhancing existing tumour biobanks in European prospective cohort studies

Tomotaka Ugai,[a,b,c,d,*] Bethany van Guelpen,[e,f] Lorelei A. Mucci,[b,c,g] and Shuji Ogino[a,b,c,h,i,j,k,**]

[a]Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, 02115, USA
[b]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, 02115, USA
[c]Cancer Epidemiology Program, Dana-Farber Harvard Cancer Center, Boston, MA, 02215, USA
[d]Division of Integrative Cancer Research, National Cancer Center Research Institute, Tsukiji, Tokyo, 104-0045, Japan
[e]Department of Diagnostics and Intervention, Oncology Unit, Umeå University, Umeå, Sweden
[f]Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden
[g]Discovery Science, American Cancer Society, Atlanta, GA, 30144, USA
[h]Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA
[i]Cancer Immunology Program, Dana-Farber Harvard Cancer Center, Boston, MA, 02215, USA
[j]Division of Nutrition, Harvard Medical School, Boston, MA, 02115, USA
[k]Institute of Science Tokyo, Tokyo, Japan

The Lancet Regional Health - Europe 2025;53: 101293

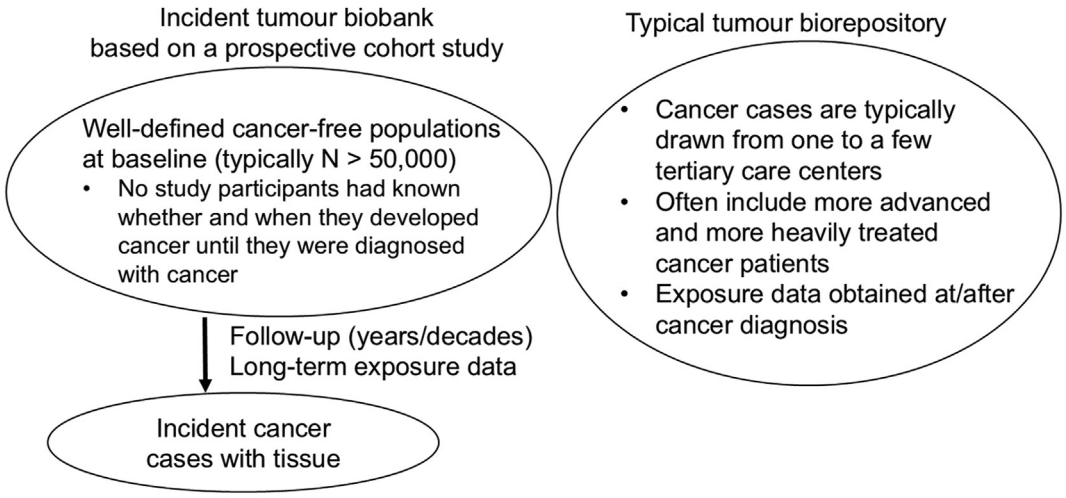
Published Online xxx https://doi.org/10.1016/j.lanepe.2025.101293

Tumour biorepositories are essential in identifying tumour molecular subtypes and biomarkers. Most tumour biorepositories include tissues from cancer patients in clinical studies and/or hospitals with which investigators are affiliated. Inherent limitations of such repositories include risk of selection bias due to convenience sampling (e.g., selection towards more advanced and more heavily treated cancer patients at tertiary care centres), lack of (or recall bias) in long-term pre-diagnostic exposure information, and inability to calculate the incidence rates of both overall cancer and tumour subtypes. Tumour tissue specimens collected from incident cancer cases nested within prospective epidemiology cohorts can address some of these challenges because of many reasons (Fig. 1).

Why are integrative analyses of long-term exposures and tumour tissue important? Cancer is a heterogeneous, multifactorial, environmental, systemic disease with different sets of cellular genetic and epigenetic alterations.[1,2] Linking long-term pre-diagnostic exposures with tumour molecular and microenvironmental characteristics can offer deeper aetiologic insights that cannot be achieved when an organ-specific cancer is considered a single disease. There is a good example in the U.S. The U.S.-based Channing/Harvard Cohorts Biorepository integrates tumour tissue samples from incident cancer cases with long-term pre-diagnostic exposure data in the Nurses' Health Study (NHS) I/II and the Health Professionals Follow-up Study (HPFS),[3,4] totalling nearly 290,000 participants followed for decades through biennial questionnaires, the U.S. National Death Index, and medical record reviews. Furthermore, this Biorepository is overlaid with additional biospecimens, including blood, urine, and stool, prospectively collected before cancer diagnosis. Utilising such resources, a novel link between long-term aspirin use and decreased incidence of PTGS2-overexpressed colorectal cancer was discovered in 2007.[5] However, no other group has validated this finding or many other findings using similar study designs.

In Europe, there are ample potential opportunities. For example, in Sweden, tumour samples collected within routine clinical diagnostics (generally in public or publicly funded health care) are stored indefinitely and can be identified for access using the Swedish Cancer Registry. Corresponding clinical data can be obtained from national quality registries and health-care records. This has been done for colorectal cancer within the Northern Sweden Health and Disease Study (NSHDS),[6] a population-based cohort comprising health- and lifestyle data and blood samples for over 140,000 participants, around half of whom have repeated sampling typically at ten-year intervals. Major molecular subtypes of colorectal cancer have thus been assessed in relation to exposures, including blood-based biomarkers, up to many years prior to colorectal cancer diagnosis.[7] Similarly, the Netherlands Cohort Study (NLCS)[8,9] has established pioneering incident colorectal cancer and renal cell tumour biobanks linked to lifestyle information at study enrolment. Tissue collection efforts have also been made within the Norfolk and Heidelberg centres of the European Prospective Investigation into Cancer and Nutrition (EPIC).[10] Fresh frozen tissue specimens are amenable to various laboratory assays; however, their collection is not part of routine clinical

*Corresponding author. Department of Pathology, Brigham and Women's Hospital 221 Longwood Ave., EBRC Room 415, Boston, MA, 02215, USA.
**Corresponding author. Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital 221 Longwood Ave., EBRC Room LM-422B, Boston, MA, 02115, USA.
E-mail addresses: tugai@bwh.harvard.edu (T. Ugai), sogino@bwh.harvard.edu (S. Ogino).



Pros	↓ Recall bias ↓ Reverse causation risk ↓ Selection bias Can directly measure incidence in overall cancer and molecular subtypes	↑ Sample size for rare cancer types ↓ Need for funding & commitments Fresh tissue collection is possible
Cons	Difficulty in studying rare cancer types Need for long-term funding & commitments Mostly limited to archival tissues	↑ Recall bias ↑ Reverse causation risk ↑ Selection bias Cannot measure incidence

Fig. 1: Comparisons of incident tumour biobanks and typical tumour biorepositories. To create an incident tumour biobank, a prospective cohort study needs to follow a large population for decades, so that enough numbers of incident cancer cases occur among the participants. Tissue specimens drawn from resected or biopsied tumour tissues are collected. Hence, it requires long-term funding and commitments of investigators. In a typical tumour biorepository, tissue specimens are usually drawn from resected or biopsied tumour tissue in one to a few hospitals, often tertiary care centres, for convenience. Pros and cons of both approaches are described.

practice. In contrast, formalin-fixed, paraffin-embedded (FFPE) tissue blocks can be more easily utilised, especially in population-based prospective cohort studies. The linkage of tumour profiling to clinical data and data from other biospecimens, including blood, urine, and stool, is warranted. Therefore, expanding initiatives for the integration of pathology and population sciences will be enormously beneficial in Europe.

How can we overcome potential challenges? To fully leverage the potential of European incident tumour biobanks, we encourage the following measures. First, there is a need for multidisciplinary collaborative teams of investigators who understand the incredible potentials of those incident tumour biobanks and dedicate their efforts to vitalising the resources. To allow adequate tissue collection that can be used both for clinical and research purposes and advance collaborations between departments of pathology and population sciences, an integrative division or department of pathology and population sciences could be created. In addition, opportunities to find mutual goals of diverse collaborators, such as conferences and symposia, and funding mechanisms to spark interdisciplinary collaborations are helpful. Trainees can often be the glue for such collaborations. Second, we need long-term secure funding and mechanisms to support the infrastructure and investigator teams. This includes support related to meeting ethico-legal requirements, in particular, handling and storage of sensitive personal data under the European Union General Data Protection Regulation (GDPR). Funding agencies and programs need to recognise the importance of long-term support and funding for rare and uniquely-valued tumour biobanks within prospective cohort studies. Because biobank-based research projects often need to be conducted across countries beyond Europe, administrative support to deal with different nation-level legislations beyond GDPR is crucial. Lastly, it is necessary to develop trans-interdisciplinary education and training programmes in Europe. It is foreseeable that integrative analyses will generate numerous novel findings to attract attention and interest of researchers with expertise in a wide variety of fields and that further integration of new research fields can lead to novel directions of interdisciplinary pathobiological population sciences.

Contributors

T.U. and S.O. wrote the original draft. T.U., B.v.G., L.A.M., and S.O. reviewed and edited the manuscript, and approved the final version.

Declaration of interests

L.A.M. reports research funding from Astra Zeneca to Harvard University; and consults and holds equity in Convergent Therapeutics. B.v.G. received a lecturer honorarium from AstraZeneca AB for an educational activity. These are unrelated to the current work. The other authors declared no conflicts of interest.

Acknowledgements

Funding: The work for this Comment was supported in part by grants from the USA National Institutes of Health (R50 CA274122 to T.U.; U01 CA167552 to L.A.M.; R01 CA248857 to S.O.), the Cancer Research UK Cancer Grand Challenge Award (C10674/A27140 to S.O.), the American Cancer Society Clinical Research Professor Award (CRP-24-1185864-01-PROF to S.O.), a grant from the Prevent Cancer Foundation (to T.U.), the Harvey V. Fineberg Cancer Prevention Fellowship (to T.U.), the Brigham and Women's Hospital Faculty Career Development Award (to T.U.), and an Investigator Initiated Grant from the American Institute for Cancer Research (to T.U.). The funders had no role in the preparation of the manuscript.

References

1 Lee HY, Song M, Stopsack KH, et al. The cancer spectrum theory. *Cancer Discov.* 2024;14(4):589–593.
2 Inamura K, Hamada T, Bullman S, Ugai T, Yachida S, Ogino S. Cancer as microenvironmental, systemic, and environmental diseases: opportunity for transdisciplinary microbiomics science. *Gut.* 2022;71(10):2107–2122.
3 Ogino S, Ugai T. The global epidemic of early-onset cancer: nature, nurture, or both? *Ann Oncol.* 2024;35(12):1071–1073.
4 Ugai T, Väyrynen JP, Ugai S, et al. Long-term marine ω-3 polyunsaturated fatty acids intake in relation to incidence of colorectal cancer subclassified by macrophage infiltrates. *Innovat Med.* 2024;2(3):100082.
5 Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. *N Engl J Med.* 2007;356(21):2131–2142.
6 Spath F, Wennberg P, Johansson R, et al. Cohort profile: the Northern Sweden Health and Disease Study (NSHDS). *Int J Epidemiol.* 2024;54(1):dyaf004.
7 Myte R, Harlid S, Sundkvist A, et al. A longitudinal study of prediagnostic metabolic biomarkers and the risk of molecular subtypes of colorectal cancer. *Sci Rep.* 2020;10(1):5336.
8 van Engeland M, Weijenberg MP, Roemen GM, et al. Effects of dietary folate and alcohol intake on promoter methylation in sporadic colorectal cancer: The Netherlands cohort study on diet and cancer. *Cancer Res.* 2003;63(12):3133–3137.
9 van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol.* 1990;43(3):285–295.
10 Le Cornet C, Walter B, Sookthai D, et al. Circulating 27-hydroxycholesterol and breast cancer tissue expression of CYP27A1, CYP7B1, LXR-beta, and ERbeta: results from the EPIC-Heidelberg cohort. *Breast Cancer Res.* 2020;22(1):23.