



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/195768/>

Version: Accepted Version

---

**Article:**

Spencer, K (2023) Identifying the unseen and unmet; using data to target blind spots in cancer care. *Journal of Cancer Policy*, 35. 100409. ISSN: 2213-5383

<https://doi.org/10.1016/j.jcpo.2023.100409>

---

© 2023 Elsevier Ltd. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

## **Identifying the unseen and unmet; using data to target blind spots in cancer care**

Dr Katie Spencer<sup>1,2,3</sup>

1. Leeds Institute of Health Sciences, University of Leeds
2. Leeds Teaching Hospitals NHS Trust, Leeds, UK
3. National Radiotherapy Dataset, NHS Digital

Despite improvements in overall cancer survival over recent decades, the survival gap, resulting from worse cancer survival in more socio-economically deprived populations, remains unchanged.(1,2) In colorectal cancer alone, if the most deprived patients in the UK had the same cancer outcomes as their affluent peers 15,000 life years would be saved each year.(3) In order to understand the causes of this gap and ultimately close it, we urgently need methodologically robust analyses of high quality data.

Previous studies have identified late stage at presentation as a contributor to the survival gap. For example, in breast and prostate cancer, this has been identified as the primary explanatory factor.(2) Variation exists, however, between studies, across jurisdictions and with cancer diagnosis. Beyond stage at diagnosis, treatment variation is associated with survival; in lung cancer, lower utilisation of surgery and radiotherapy has been found to be associated with poorer survival.(4) Some variation maybe justified, as a result of observable confounders such as co-morbidity, or indeed, unobserved confounders related to the social determinants of health.(2) The challenge lies in understanding the extent to which these different factors contribute to identified inequalities. Approaches aiming to improve early diagnosis, for example through cancer screening or public symptom awareness campaigns, are dramatically different to those that might address geographical or provider level variation in treatment access, such as targeted resource investment or wider modifications to health service commissioning. Work in this area continues but the challenges faced by researchers using these data have increased over the last decade. Justifiable concerns for patient confidentiality combined with, at times, overly cautious interpretation of data protection legislation in some jurisdictions, make data access processes inefficient, costly and time consuming with linkage to other data sources frequently impossible. Despite increasingly being surrounded by data, we have not managed to harness this to understand and improve equity in cancer outcomes.

Multiple mechanisms are required to improve this, first and foremost, communication with the wider public about the benefits of using health data for research has been sorely limited with the resulting loss of trust leading to individuals opting out of sharing their data. The comprehensive analyses required to understand and address inequalities will become impossible if opt-outs rise beyond a critical level; communication with patients and the public to articulate the benefits of using their data to build equitable, effective and efficient cancer services, must be central to future progress, not an after thought.

International cancer registries, enabled by strong public engagement, can deliver the comprehensive population coverage required to deliver these analyses. Registration data must, however, be supplemented with linked baseline and treatment data to support elucidation of the complex reasons behind the survival gap. Defining the necessary data items and ensuring inter-operability, ideally across jurisdictions, will require inter-disciplinary collaboration and public engagement. The data required to consider access to radiotherapy, chemotherapy or surgery differ widely, with these differences only extending when the focus of analysis expands, for example, to incorporate health economics or bioinformatics. This multi-stakeholder engagement is critical to ensuring the priorities

of patients and methodologists are clearly understood and translated into the development of inter-operable datasets that can support a broad range of analyses.

To date investment in health services research in cancer has, however, lagged far behind laboratory and pharmaceutical investment. In order to deliver the necessary inter-disciplinary perspective and robust analyses, strategic investment is required in the health services research workforce. Routine data are messy and understanding causal relationships is challenging. By building research capacity and capability, including skills in causal inference, we can maximise the benefit of these data. Combining the necessary high-fidelity data with methodological rigour will not only enable us to understand and address inequalities, but also to fulfil the promise of real-world evidence in efficiently assessing treatment efficacy and value whilst addressing the challenges of bias inherent in such analyses.

Beyond the crucial role of the research workforce there is an urgent need to develop the infrastructure and processes required to provide a balance between efficient, appropriate data access on the one hand, and data security that delivers confidence confidentiality will be maintained on the other. Realistic appraisals of the risks posed by anonymised data use are required, trusted digital research environments developed and data access enabled, where necessary supported by contractual agreements with penalties (e.g. loss of future data access and/or fines) for breaches. In Europe the recently proposed European Health Data Space has the potential to improve the accessibility and use of healthcare data for the benefit of patients with cancer across Europe, although greater detail is urgently required on its implementation.(5) Parallel infrastructure is under development in England and Wales.(6,7) The current progress undoubtedly requires acceleration, however, this technical infrastructure must also be supported by data management capacity in data controllers. Without high fidelity datasets there is a risk that multiple 'versions of the truth' exist, undermining confidence in the delivered analyses. Systematic sharing of analytical code can then deliver greater research efficiency and reproducibility.(8) This analytical code sharing is only of value though if data are available; not all routine healthcare datasets are held by public bodies. Ensuring fair, efficient and affordable access to inter-operable data held by commercial organisations, for research and service development, will be critical to delivering sustainability for the wider data analysis ecosystem.

In all of these areas pockets of excellence exist globally but we can, and must, do better. Success will not only reduce inequities but also support wider health services research, increasing the effectiveness and efficiency of cancer care, closing the gap and improving outcomes for all. It is time to target our blind spots and deliver on the promise of routine cancer data for the benefit of all.

## References:

1. Rachet B, Ellis L, Maringe C, Chu T, Nur U, Quaresma M, et al. Socioeconomic inequalities in cancer survival in England after the NHS cancer plan. *Br J Cancer*. 2010 Aug;103(4):446–53.
2. Afshar N, English DR, Milne RL. Factors Explaining Socio-Economic Inequalities in Cancer Survival: A Systematic Review. *Cancer Control*. 2021 Jan 1;28:10732748211011956.
3. Syriopoulou E, Morris E, Finan PJ, Lambert PC, Rutherford MJ. Understanding the impact of socioeconomic differences in colorectal cancer survival: potential gain in life-years. *Br J Cancer*. 2019 May;120(11):1052–8.

4. Tataru D, Spencer K, Bates A, Wieczorek A, Jack RH, Peake MD, et al. Variation in geographical treatment intensity affects survival of non-small cell lung cancer patients in England. *Cancer Epidemiology*. 2018 Dec;57:13–23.
5. European Health Data Space [Internet]. [cited 2022 Dec 16]. Available from: [https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space\\_en](https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en)
6. Trusted Research Environment service for England [Internet]. NHS Digital. [cited 2022 Dec 16]. Available from: <https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england>
7. SAIL DATABANK [Internet]. SAIL Databank. [cited 2022 Dec 16]. Available from: <https://saildatabank.com/>
8. Goldacre B, Morley J. A review commissioned by the Secretary of State for Health and Social Care. Better, broader, safer: using health data for research and analysis. 2022 Apr [cited 2022 Dec 16]; Available from: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1067053/goldacre-review-using-health-data-for-research-and-analysis.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1067053/goldacre-review-using-health-data-for-research-and-analysis.pdf)