The full reference dataset with annotations reported by Ammons et al. was downloaded from Zenodo1,2. All data analysis steps were preformed using R version 4.3.2 and the *tidyverse* suit of R packages v 2.0.03,4. We benchmarked three deconvolution tools (*SCDC*5, *DWLS*6 and *MuSiC*7) on this reference dataset by generating 100 pseudobulk samples by randomly selecting 10K cells from the reference dataset. Each pseudobulk sample contained cells representing cell types from “celltype.l1” in random proportions. Only genes shared by the bulk data and single-cell reference were included in the pseudobulk samples. Length normalization was not performed on the pseudobulk matrix because it was generated using cells processed with a Chromium Next GEM Single Cell 3ʹ v3.1 kit.

The length of the merged exons of isoforms of each gene was calculated from CanFam3.1 database version 100 using GTFtools8 and used to TPM normalized the bulk data. Only genes with ensembl IDs present in the the raw count data and the annotation file were included. Gene symbols are used in the final bulk matrix to match the single-cell reference data. Duplicated gene symbols were removed.

*MuSiC* v 0.3.0 and *DWLS* v 0.4 were run using the wrapper R package *omnideconv* v 0.1.0 using the default options9. *SCDC* v 0.0.0.9000 was run with a qc-threshold of 0.50.

1. Ammons, D. T. *et al.* Single-cell RNA sequencing reveals the cellular and molecular heterogeneity of treatment-naïve primary osteosarcoma in dogs. *Commun Biol* **7**, 496 (2024).

2. Dylan Ammons. Canine osteosarcoma single-cell RNA sequencing reference dataset: analysis code and processed data for publication. (2024) doi:10.5281/ZENODO.10891255.

3. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing (2021).

4. Wickham, H. *et al.* Welcome to the Tidyverse. *JOSS* **4**, 1686 (2019).

5. Dong, M. *et al.* SCDC: bulk gene expression deconvolution by multiple single-cell RNA sequencing references. *Brief Bioinform* doi:10.1093/bib/bbz166.

6. Tsoucas, D. *et al.* Accurate estimation of cell-type composition from gene expression data. *Nat Commun* **10**, 2975 (2019).

7. Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nature Communications* **10**, 1–9 (2019).

8. Li, H.-D., Lin, C.-X. & Zheng, J. GTFtools: a software package for analyzing various features of gene models. *Bioinformatics* **38**, 4806–4808 (2022).

9. Dietrich, A. *et al.* Benchmarking second-generation methods for cell-type deconvolution of transcriptomic data. Preprint at https://doi.org/10.1101/2024.06.10.598226 (2024).