Decosus: An R Framework for Universal Integration of Cell Proportion Estimation Methods

Chinedu A. Anene1,2*, Emma Taggart3, Catherine A. Harwood4,5, Daniel J. Pennington3 and Jun Wang1

1Centre for Cancer Genomics and Computational Biology, Barts Cancer Institute, Queen Mary University of London, London, United Kingdom, 2Centre for Cancer Biology and Therapy, School of Applied Science, London South Bank University, London, United Kingdom, 3Centre for Immunobiology, Barts and the London School of Medicine, Bizzar Institute, Queen Mary University of London, London, United Kingdom, 4Centre for Cell Biology and Cutaneous Research, Barts and The London School of Medicine and Dentistry, Bizzar Institute, Queen Mary University of London, London, United Kingdom, 5Department of Dermatology, The Royal London Hospital, Barts Health NHS Trust, London, United Kingdom

The assessment of the cellular heterogeneity and abundance in bulk tissue samples is essential for characterising cellular and organismal states. Computational approaches to estimate cellular abundance from bulk RNA-Seq datasets have variable performances, often requiring benchmarking matrices to select the best performing methods for individual studies. However, such benchmarking investigations are difficult to perform and assess in typical applications because of the absence of gold standard/ground-truth cellular measurements. Here we describe Decosus, an R package that integrates seven methods and signatures for deconvoluting cell types from gene expression profiles (GEP). Benchmark analysis on a range of datasets with ground-truth measurements revealed that our integrated estimates consistently exhibited stable performances across datasets than individual methods and signatures. We further applied Decosus to characterise the immune compartment of skin samples in different settings, confirming the well-established Th1 and Th2 polarisation in psoriasis and atopic dermatitis, respectively. Secondly, we revealed immune system-related UV-induced changes in sun-exposed skin. Furthermore, a significant motivation in the design of Decosus is flexibility and the ability for the user to include new gene signatures, algorithms, and integration methods at run time.

Keywords: cell deconvolution, R package, method integration, gene expression, immuno-biology

INTRODUCTION

Gene expression quantification is indispensable for the interrogation of cellular and organismal states. However, bulk tissue samples of interest in clinical research have considerable cellular heterogeneity that standard methods (Microarray and RNA-Seq) cannot decipher. Although single-cell technologies have been developed to uncover the cellular heterogeneity within cell populations, they have a range of limitations (e.g., time, tissue types, dropouts, technical noise, and cost), making large-scale or clinical applications impossible. Since bulk analysis methods report average expression levels, it is often challenging to disentangle changes in cell-type composition from fundamental differences in states. To address this, several computational tools (so-called deconvolution methods) have been developed to estimate cell-type composition...
within bulk expression data. These utilise various models, including least squares regression (Abbas et al., 2009), constrained least squares regression (Li et al., 2016; Racle et al., 2017; Finotello and Trajanoski, 2018), quadratic programming (Gong et al., 2011; Gong and Szustakowski, 2013; Zhong et al., 2013), support vector regression (Newman et al., 2015), the geometric mean of marker gene expression (Becht et al., 2016) and single-sample gene set enrichment analysis (ssGSEA) (Aran et al., 2017), extensively reviewed in (Sturm et al., 2019).

Unfortunately, individual tools have limitations that affect their effective utilisation. Critically, due to the differences in the underlying statistical assumptions and marker gene signatures, the approaches often produce different results in a data type-dependent manner (Jiménez-Sánchez et al., 2019; Sturm et al., 2019). Thus, there is a need for a consensus approach that can combine and integrate these methods into a single robust output.

Recent benchmarking studies (Jiménez-Sánchez et al., 2019; Sturm et al., 2019) have provided robust frameworks to systematically integrate the outputs of these methods into a single unified deconvolution solution, thereby reducing the limitations of individual methods. While these approaches are helpful, significant challenges remain. Jiménez-Sánchez et al., restricted their implementation to a set of common cancers (Jiménez-Sánchez et al., 2019), while Sturm et al., focused on method performance on key immune cell types (Sturm et al., 2019). Thus, neither tool can be applied universally, especially for poorly studied cancers, and non-cancer conditions or deconvolution of non-immune cells. We developed a deconvolution integration tool (Decosus, available as an R package) without tissue type restrictions, allowing for universal application of the integration framework. A further advantage of Decosus is the ability to generate both relative (allowing sample-to-sample comparison only) and absolute (also allowing within-sample comparison of one cell type to another) estimations where possible. We demonstrate the utility of Decosus through reference datasets and use cases, including cancer and non-cancer datasets. Decosus generates a validated and robust decomposition of cell composition in various tissue types while allowing for flexible integration of new methods and signatures as they become available (github.com/caanene1/Decosus).

**MATERIALS AND METHODS**

**Deconvolution Method**

We used common cell deconvolution methods and signatures to estimate cell composition from bulk RNA-Seq datasets individually. The methods include EPIC (Racle et al., 2017), MCP-counter (Becht et al., 2016), quanTISeq (Finotello et al., 2019), and xCell (Aran et al., 2017). The estimates for the four methods were generated using the corresponding R implementations: MCPCounter (v1.2.0), EPIC (v1.1.5), xCell (v1.1.0), and quanTISeq (github.com/ici-bi-lab/quanTIseq). In addition to these, we used the following gene signatures: Danaher, Davoli, Rooney (Bindea et al., 2013; Danaher et al., 2017; Davoli et al., 2017), and averaged the gene expressions to provide measures of cellular abundance for each signature. Additionally, we provide optional gene signatures from the CellMarker database (Zhang et al., 2019) and use this in the analyses featured here.

### Decosus Framework

Decosus is an R package that flexibly integrates the estimates of cell compositions from deconvolution methods and cell signatures. We identified cell types for which estimates exist in at least two of the seven default methods and signatures (Supplementary S1 CSV). This mapping was similar to that of a previously published method, ConsensusMED, where cells were exhaustively mapped to a controlled vocabulary of cell types (Jiménez-Sánchez et al., 2019). However, in Decosus, the user can expand the consensus through the optional arguments for new signatures and mappings (https://github.com/caanene1/Decosus), ensuring that Decosus can produce results relevant to the most up-to-date cell signatures and allowing the addition of rarer cell types. For instance, licensing restrictions did not allow us to include CIBERSORT into the package, but a user can still integrate CIBERSORT if you obtain the output or source code from the author’s website.

After generating the estimates for the individual methods, we average the values for the source cells to create a single estimate for the cell types in the controlled vocabulary (Supplementary S1 CSV). The assumption behind this approach is that it limits any one method from dominating the estimates, thereby ensuring the consensus is closer to the best measure or is the best performing measure (see results section). We also included optional arguments in the function call to specify whether the data is to be scaled or not and the aggregation mode (i.e., mean (default) or geometric mean). Furthermore, two outputs are provided; 1) relative, the default using all methods, and 2) absolute, which is derived by limiting the methods used to those that can be considered absolute cell compositions (EPIC and quanTISeq, reviewed previously by Sturm et al., 2019). This option offers a less comprehensive selection of consensus cells than the relative output but may be helpful in analyses requiring cell-cell comparison, which is not permitted when using all seven algorithms due to the methods involved (see Table 1 and Sturm et al., 2019). The Decosus R package can be obtained through GitHub (https://github.com/caanene1/Decosus).

### Table 1: Default deconvolution method and signatures included in Decosus.

| Signature/Method | Type | Comparison |
|------------------|------|------------|
| XCell            | Algorithm | Across Samples |
| MCP-counter      | Algorithm | Across Samples |
| quanTISeq        | Algorithm | Across Samples & Between Cells |
| EPIC             | Signature | Across Samples |
| Danaher          | Signature | Across Samples |
| Davoli           | Signature | Across Samples |
| Rooney           | Signature | Across Samples |

*Default deconvolution method and signatures included in Decosus.*
Datasets for Benchmarking
To benchmark the Decosus framework, we obtained pre-processed bulk RNA-seq and FACS data from Sturm et al. (2019) through their Github repository (https://github.com/icbi-lab/immune_deconvolution_benchmark/releases/download/v1.0.0-rcl/data.tar.gz). This data includes eight healthy PBMC samples (Hoek data) (Finotello et al., 2019), four metastatic melanoma samples (Racle data) (Racle et al., 2017) and three ovarian cancer ascites samples (Schelker data, x2 replicates) (Schelker et al., 2017). We obtained 20 PBMC samples analysed by microarray (gene-level values) and flow-cytometry from the CIBERSORT web portal (Cibersort data) (Newman et al., 2015). Additionally, we obtained the two (SDY311, n = 76; SDY420 ref, and n = 105) pre-processed bulk RNA-seq and FACS data from the Immport study (Alpert et al., 2019) through the xCell Github repository (https://github.com/dviraran/xCell/tree/master/vignettes).

To evaluate the performance of Decosus in real-world clinical contexts, we interrogated two additional RNA-Seq datasets. The first dataset contained skin samples of atopic dermatitis (n = 54) and psoriasis (n = 55) (GSE121212) (Tsoi et al., 2019). The second dataset was normal skin samples (sun-exposed and non-sun exposed) samples from the GTEX project (n = 1879) (Lonsdale et al., 2013).

All RNA-Seq gene expression values are expressed as transcript per million (TPM).

Datasets for Simulation Studies
To evaluate how well Decosus estimates known cell compositions compared to individual methods, we simulated cell mixtures from single-cell RNA-Seq datasets. Specifically, we extracted the RNA expression values for select samples from the Panglaodb data portal (Franzén et al., 2019), including SRA701877, SRS3279685, SRA713577, SRS3363004, SRA716608, SRS3391633, SRA779509, SRS3805246, SRA878024, and SRS4660846.

Next, we generated expression profiles as below:

1. Given a dataset of annotated n single cell types, assign random fractions to each cell type (the fractions sum to 100 and integers) (data 1).
2. Generate an expression matrix of single cells with 100 columns by randomly selecting the corresponding fraction of the available samples for the selected cell types to be included in the matrix. Here, we used a random selection to introduce noise like the variation in real datasets (data 2). We allow sampling with replacement if the cell type-specific fraction is bigger than the available single cells.
3. Finally, generate a simulated expression profile by adding the expression values across the rows of data 2 and use data 1 as the ground truth.

We repeat the process 500 times (per data source) with different fractions, samples, and cell types.

RESULTS
Overview of the Decosus Integration Framework
Our framework integrates seven deconvolution algorithms and cell signatures into consensus estimates of cell composition (Table 1). We do this by selecting and averaging the shared cell types across the tools (Supplementary S1 CSV). It requires a gene expression dataset and a set of optional parameters under a single R function (set to reasonable defaults, see github.com/caanene1/Decosus). The R implementation of the framework allows for flexible inclusion of new algorithms or signatures at run time (see https://github.com/caanene1/Decosus). When available, the final output has two tables representing relative and absolute consensus estimates (see Materials and Methods).
Analysis of Simulated Cell Proportions Demonstrate the Stability of Decosus

Cell proportion deconvolution methods produce different results in a data type-dependent manner. Decosus combines and integrates these methods into a single stable consensus value to reduce the data-dependent differences in performance. To assess how well the framework achieves this aim, we first evaluated its ability to decompose known cell proportions using simulated gene expression profiles (see Methods). We simulated large sets of bulk-expression profiles (n = 2,500) with specific cell proportions from multiple single-cell RNA-Seq datasets (source, n = 5). We

FIGURE 2 | Decosus identifies known cell fractions. Bar plots of the Pearson’s R of the correlation between computer-generated cell fractions and flow cytometry fractions. The higher the R, the more similar is the computational estimates and the gold standard. Within the plots colours represent the method of estimating cell fractions from bulk RNA-Seq datasets. MCP = MCPcounter.
used multiple sources to reflect data-specific differences in real applications. Interrogating the correlation coefficient between the estimates and the expected cell proportions across a range of cell types revealed surprisingly stable estimates of cell proportions (Figure 1). Decosus had the 4th lowest interquartile range (IQR = 0.31) compared to xCell (IQR = 0.39), MCPcounter (IQR = 0.47), quanTISeq (IQR = 0.50), and Danaher (IQR = 0.64) (Figure 1A). Although Rooney (IQR = 0.09), Davoli (IQR = 0.22), and EPIC (IQR = 0.29) were lower than Decosus, this is potentially due to the small number of cell types they covered in the simulation (Davoli = 3, EPIC = 3, and Rooney = 2), compared to the other methods (Decosus = 6, xCell = 6, MCPcounter = 5, quanTISeq = 4, and Danaher = 4). Indeed, normalising for the number of cell types revealed Decosus has the second-lowest IQR (Supplementary Figure S1). The stability of Decosus estimates is due to the robustness of averaging multiple signatures and methods (see Methods). As expected, there is no difference between median R values across the methods (Kruskal–Wallis, p = 0.6, Figure 1B), suggesting that Decosus increases the stability of the estimates without reducing the average performance expected from individual tools.

**Benchmark on Known Cell Proportions Demonstrate the Utility of Decosus**

We applied the Decosus framework to six benchmark datasets (see Methods) and compared the estimated proportions with their corresponding flow cytometry fractions used as ground-truth. We also interrogated the distribution of the performance statistic across the individual methods and datasets. To ensure we can perform meaningful comparative analysis, we restricted our analysis to six cell types (natural killer cells, dendritic cells, monocytes, CD4+ T-cells, CD8+ T-cells and B-cells) present in two or more ground truth datasets. Our consensus estimates obtained a high correlation with the corresponding gold standards (Figure 2). Decosus and many existing methods do not have these two cell types, thus could not handle such cases. To this end, we collected 54 AD and 55 PSO samples from a publicly available dataset (GSE121212) (Tsai et al., 2019) and applied our framework to interrogate the enriched immune profiles. We observed a significantly higher Th1 cell signature in PSO samples than AD samples and the reverse for Th2 (Figure 3), aligning with what is widely recognised in the literature (Brunner et al., 2017; Albanesi, 2019). Thus, immune cell composition estimated from bulk RNA expression profiles from PSO and AD skin samples should enrich for Th1 and Th2 cell signatures, respectively. Note that many existing methods do not have these two cell types, thus could not handle such cases. To this end, we collected 54 AD and 55 PSO samples from a publicly available dataset (GSE121212) (Tsai et al., 2019) and applied our framework to interrogate the enriched immune profiles. We observed a significantly higher Th1 cell signature in PSO samples than AD samples and the reverse for Th2 (Figure 3), aligning with what is widely recognised in the literature (Brunner et al., 2017; Albanesi, 2019). The results also showed expansions in other cell types known to infiltrate each lesion, such as basophils in AD (Mashiko et al., 2017), macrophages and neutrophils in PSO (Lowes et al., 2014) (Figure 3). These results indicate that our framework can provide robust estimates of cell proportions in non-cancer bulk tissue samples, like precursor lesions.

**Decosus Enables the Identification of UV Mediated Immune Reprogramming in the Skin**

Finally, we utilised transcriptomic data from the GTEx project of non-sun exposed and sun-exposed skin to evaluate the
impact of ultraviolet radiation (UV) on skin immune profiles. We assessed the difference between the immune cell estimates of the sample groups and visualised the fold change for significant (T-test, $p < 0.05$) cell types. We found that UV exposure significantly enriched several immune cell types in the skin, including monocytes, dendritic cells, and macrophages (Figure 4). Interestingly, CD4$^+$ and CD8$^+$ T cells were depleted in sun-exposed skin (Figure 4). These observations are consistent with previous studies showing that UV exposure inhibits the expansion of these T cell subtypes while increasing innate immune cells (Rana et al., 2008; Gläser et al., 2009). Reprogramming of T cell composition is consistent with the idea that impaired immune function through UV damage plays a role in skin cancers (Freeman et al., 2014; Slater and Googe, 2016). Indeed, active research programmes in our group are using Decosus to help characterise the immunological factors underlying the progression of actinic keratoses (sun-damaged skin, pre-malignant lesions) to squamous cell carcinomas.
DISCUSSION

Estimating the cellular compositions in bulk RNA-Seq samples has been addressed by various algorithms and gene signatures. These approaches have different assumptions and strengths, frequently resulting in varied performances across different datasets. Unlike the ConsensusTME, which generated new consensus gene signatures (Jiménez-Sánchez et al., 2019), we solved this problem by first calculating the cell estimations individually for each method (see Methods) before taking the average value for the same cell in each method. Here, a future update may include additional ways to generate consensus estimates such as geometric mean, trimmed and weighted averages, particularly as validation datasets become available. Further advantages to our tool are the ability to apply Decosus to any tissue type or disease compared to previous efforts to create consensus cell composition estimates, which focused on tumor tissues (Jiménez-Sánchez et al., 2019; Sturm et al., 2019). We also incorporated the option to use absolute estimates, which is crucial for applications requiring cell-cell comparisons.

Benchmarking on the PBMC and cancer datasets showed that although each method ranked highly in at least one dataset, we saw highly varied correlations across datasets. This performance issue is expected but impossible to identify in actual use cases because of the differences in the statistical assumptions and gene signatures associated with each method. Our consensus estimates reduce this unpredictable behaviour by averaging out the poorly performing methods. Indeed, we adequately identified the expected Th1 cells enriched in PSO and the Th2 cells in AD. It is worth noting that many existing methods, such as Cibersort, EPIC, and ConsensusTME, could not resolve such a case due to their limited coverage of diseases/conditions or cell types (see Supplementary S2 CSV), further highlighting the versatility of our method.

An important limitation of the Decosus framework is that it represents the aggregate performance of the contributing methods. Thus, if they have universally poor performance for a given case, then Decosus will have a corresponding poor performance. For instance, all the approaches, including Decosus, performed poorly in decomposing NK cells and Monocytes from the SDY311 dataset (Figure 2). However, Decosus is stable for most cases compared to the individual methods.

To allow for easy incorporation of Decosus into new and existing workflows, we implemented an object-oriented system in R, allowing the user to add, retrieve and evaluate individual methods (https://github.com/caanene1/Decosus). The full output of Decosus provides consensus estimates when available and offers unified interface for the procedures. Although, Decosus is implemented and valid for human data, the framework can easily be expanded with new methods and signatures, including for other species at run time, as demonstrated in the analysis of Th1 and Th2 cells in skin samples. Moreover, one of the future directions of Decosus is to create a flexible function within our method that allows users to input associated weights of individual methods and additive equations, to facilitate users to infer the most accurate estimates of cellular compositions in their biological settings. However, this will require very large validation datasets to derive weights accurately for the biological setting of interest.
DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualisation, CA, ET, CH, DP and JW; Software, CA and ET; Investigation, CA, ET, CH, DP and JW; Writing—Original Draft, CA and ET; Writing—Review & Editing, CA, ET, CH, DP and JW; Supervision, CH, DP, and JW.

FUNDING

This study was supported by the Barts Charity project grant (reference MRDU0003), awarded to CH, DP, and JW.

REFERENCES

Abbas, A. R., Wolseleg, K., Sesahsayee, D., Modrusan, Z., and Clark, H. F. (2009). Deconvolution of Blood Microarray Data Identifies Cellular Activation Patterns in Systemic Lupus Erythematosus. PLoS One 4, e6998. doi:10.1371/journal.pone.0006998

Albanesi, C. (2019). “Immunology of Psoriasis,” in Clinical Immunology (Elsevier). 871–878. doi:10.1016/b978-0-7020-6896-6.00064-8

Alpert, A., Pickman, Y., Leipoldt, M., Rosenberg-Hasson, Y., Ji, X., Gaujoux, R., et al. (2019). A Clinically Meaningful Metric of Immune Age Derived from High-Dimensional Longitudinal Monitoring. Nat. Med. 25, 487–495. doi:10.1038/s41591-019-0381-y

Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. Genome Biol. 18, 220–314. doi:10.1186/s13059-017-1349-1

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. Genome Biol. 17, 218–220. doi:10.1186/s13059-016-1070-5

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. Immunity 39, 782–795. doi:10.1016/j.immuni.2013.10.003

Brunner, P. M., Gutman-Yassky, E., and Leung, D. Y. M. (2017). The Immunology of Atopic Dermatitis and its Reversibility with Broad-Spectrum and Targeted Therapies. J. Allergy Clin. Immunol. 139, S65–S76. doi:10.1016/j.jaci.2017.01.011

Danaher, P., Warren, S., Dennis, L., D’Amico, L., White, A., Disis, M. L., et al. (2017). Gene Expression Markers of Tumor Infiltrating Leukocytes. J. Immunother. Cancer 5, 18–15. doi:10.1186/s40425-017-0215-8

Davoli, T., Uno, H., Wooten, E. C., and Elledge, S. J. (2017). Tumor Aneuploidy Correlates with Markers of Immune Evasion and with Reduced Response to Immunotherapy. Science 355, eaaf8399. doi:10.1126/science.aaf8399

Finotello, F., and Trajanoski, Z. (2018). Quantifying Tumor-Infiltrating Immune Cells from Transcriptomics Data. Cancer Immunol. Immunother. 67, 1031–1040. doi:10.1007/s00262-018-2150-z

Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., et al. (2019). Molecular and Pharmacological Modulators of the Tumor Immune Contexture Revealed by Deconvolution of RNA-Seq Data. Genome Med. 11, 34–20. doi:10.1186/s13059-019-0638-6

ACKNOWLEDGMENTS

The authors acknowledge support from Cancer Research United Kingdom Centre of Excellence Award to Barts Cancer Centre (reference C16420/A18066). JW acknowledges support from the Academy of Medical Sciences Springboard Award (SBF0031025), and Cancer Research United Kingdom (C355/A26619) and FC AECC and AIRC under the Accelerator Award Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.802838/full#supplementary-material

Supplementary Figure 1 | Bar plot of inter-method variability in Pearson's R of correlation between decomposed and simulated cell fractions measured by IQR normalised by the number of cell types covered for each method in the simulation. MCP = MOPcounter.
Rana, S., Byrne, S. N., MacDonald, L. J., Chan, C. Y.-Y., and Halliday, G. M. (2008). Ultraviolet B Suppresses Immunity by Inhibiting Effector and Memory T Cells. *Am. J. Pathol.* 172, 993–1004. doi:10.2353/ajpath.2008.070517

Schelker, M., Feau, S., Du, J., Ranu, N., Klipp, E., MacBeath, G., et al. (2017). Estimation of Immune Cell Content in Tumour Tissue Using Single-Cell RNA-Seq Data. *Nat. Commun.* 8, 2032–2112. doi:10.1038/s41467-017-02289-3

Slater, N. A., and Googe, P. B. (2016). PD-L1 Expression in Cutaneous Squamous Cell Carcinoma Correlates with Risk of Metastasis. *J. Cutan. Pathol.* 43, 663–670. doi:10.1111/cup.12728

Sturm, G., Finotello, F., Petitprez, F., Zhang, J. D., Baumbach, J., Fridman, W. H., et al. (2019). Comprehensive Evaluation of Transcriptome-Based Cell-type Quantification Methods for Immuno-Oncology. *Bioinformatics* 35, i436–i445. doi:10.1093/bioinformatics/btz363

Tsoi, L. C., Rodriguez, E., Degenhardt, F., Baurecht, H., Wehkamp, U., Volks, N., et al. (2019). Atopic Dermatitis Is an IL-13-Dominant Disease with Greater Molecular Heterogeneity Compared to Psoriasis. *J. Invest. Dermatol.* 139, 1480–1489. doi:10.1016/j.jid.2018.12.018

Zhang, X., Lan, Y., Xu, J., Quan, F., Zhao, E., Deng, C., et al. (2019). CellMarker: a Manually Curated Resource of Cell Markers in Human and Mouse. *Nucleic Acids Res.* 47, D721–D728. doi:10.1093/nar/gky900

Zhong, Y., Wan, Y. W., Pang, K., Chow, L. M., and Liu, Z. (2013). Digital Sorting of Complex Tissues for Cell Type-specific Gene Expression Profiles. *BMC Bioinformatics* 14, 89–10. doi:10.1186/1471-2105-14-89

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Anene, Taggart, Harwood, Pennington and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.