nature communications

Article

https://doi.org/10.1038/s41467-023-37353-8

Pan-cancer classification of single cells in the tumour microenvironment

Received: 6 July 2022

Accepted: 10 March 2023

Published online: 23 March 2023

Check for updates

Ido Nofech-Mozes[1,2], David Soave[1,3], Philip Awadalla[1,2,4,5] ✉ & Sagi Abelson[1,2,5] ✉

Single-cell RNA sequencing can reveal valuable insights into cellular heterogeneity within tumour microenvironments (TMEs), paving the way for a deep understanding of cellular mechanisms contributing to cancer. However, high heterogeneity among the same cancer types and low transcriptomic variation in immune cell subsets present challenges for accurate, high-resolution confirmation of cells' identities. Here we present scATOMIC; a modular annotation tool for malignant and non-malignant cells. We trained scATOMIC on >300,000 cancer, immune, and stromal cells defining a pan-cancer reference across 19 common cancers and employ a hierarchical approach, outperforming current classification methods. We extensively confirm scATOMIC's accuracy on 225 tumour biopsies encompassing >350,000 cancer and a variety of TME cells. Lastly, we demonstrate scATOMIC's practical significance to accurately subset breast cancers into clinically relevant subtypes and predict tumours' primary origin across metastatic cancers. Our approach represents a broadly applicable strategy to analyse multicellular cancer TMEs.

Tumour microenvironments (TMEs) are highly complex. Various immune and stromal cells within the TME interact with cancer cells to regulate processes such as angiogenesis, tumour proliferation, invasion, and metastasis, as well as mediate mechanisms of therapeutic resistance[1–4]. Single-cell RNA sequencing (scRNA-seq) techniques are explicitly suitable to disentangle complex systems as they provide transcriptome information for every cell within a sample, enabling the study of subtle transcriptomic changes reflecting different cell types and their functional states[5].

Cell-type annotation is arguably the most critical step to derive biological insight from scRNA-seq experiments and can be performed manually or using automatic classifiers[6,7]. Manual annotation is often unfavourable as it is subjective to user definition of non-parametric clustering of cells, conducted under the assumption that all the cells within a defined cluster are identical, and depends on pre-existing knowledge of canonical genes. Although expression of canonical markers has been used to characterise some cell types, definitive markers are not always available[8]. Moreover, due to their relatively low number and the possibility of incomplete detection due to technical variation, the sole use of canonical gene expression is not ideal.

Given these limitations, there has been a shift towards automatic methods for cell classification, with over 100 classifiers described in a recent census of available scRNA tools[9]. To date, most automated annotators are focused on classification of blood or subsets of cells from other specialised tissues, thus having limited capabilities in deciphering complex TMEs across diverse human cancers. Indeed, using single-cell transcriptomics to predict cancer types and differentiate between cancer and related normal tissue cells while also classifying the large number of immune cells and stroma, is not a straightforward task[10]. In the context of TMEs, cell type predictions are challenged by high inter-patient tumour cell heterogeneity among cancers of the same tissue[11–13] and low transcriptomic variation among related, yet different specialised immune cells[14]. Since cancer samples tend to cluster by patient[11–13] and transcriptional variation is often driven by genomic instability, existing cell type classification methods

[1]Ontario Institute for Cancer Research, Toronto, ON, Canada. [2]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada. [3]Department of Mathematics, Wilfrid Laurier University, Waterloo, ON, Canada. [4]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada. [5]These authors jointly supervised this work: Philip Awadalla, Sagi Abelson. ✉e-mail: philip.awadalla@oicr.on.ca; sabelson@oicr.on.ca

which rely on distance correlations to a reference[15–17] are expected to fail.

The current standard for the identification of malignant cells in scRNA-seq data relies on copy number variation (CNV) inference methods[18,19]. Nevertheless, these methods are incapable of providing definitive information concerning the cancer's tissue of origin. Furthermore, CNV inference necessitates the presence of genetically unstable cells, and its accuracy may suffer when lacking a sizeable, distinctive normal cell reference within the sequenced specimen. Solely relying on the presence of inferred CNVs to annotate malignant cells may lead to false negatives or undefined cells in cancers with minimal genomic structural variation or nearly diploid genomes. Thus, a limitation in scRNA-seq analysis of tumour ecosystems is that there is no universal method for effective, detailed classification of heterogenous non-malignant TME cell types and subtypes, and cancer cells.

Clearly, a fully automated, pan cancer classification scheme that can easily be updated to capture additional subsets of normal cells and clinically relevant molecular subtypes of cancer, holds promise towards a better understanding of cancer ontogenies and the molecular interaction of diverse tumour tissues with their microenvironments.

In this work, we present **s**ingle **c**ell **a**nnotation of **t**um**o**ur **m**icroenvironments **i**n pan-**c**ancer settings (scATOMIC), a comprehensive, pan cancer, TME cell type classifier. We devise a structured scheme that uses hierarchically organized models and elimination processes, reducing the transcriptomic complexity of the TME multi-cellular system to improve cell classification.

Results

An overview of scATOMIC

We postulated that the sheer number of publicly available single-cell transcriptomic datasets will enable the development of a highly accurate and thorough classifier for cancer, blood, and stromal cells. To define a pan cancer reference, we interrogated cancer patient data, augmented by two additional comprehensive data sources containing transcriptomic-independent confirmation of cell identities. These include scRNA-seq of cancer cell lines representing 19 common cancer types[20] and a CITE-seq dataset (proteomics and transcriptomics) of diverse peripheral blood cells[16]. scRNA-seq of stromal cells was gathered from several tumour and normal tissue sources[21–28] (Supplementary Data 1, 2). Overall, 301,662 cells were included in the training reference dataset of scATOMIC.

Obtaining an accurate set of differentiating features is critical to successful classification. Nonetheless, significant differentially expressed genes (DEG) concerning non-malignant cell types are often expressed in other related cells that are functionally distinct[14,29,30] (Supplementary Fig. 1). On the other hand, inter-patient heterogeneity among malignant cells has been repeatedly observed with different patients forming unique clusters[11–13] (Supplementary Fig. 2). To improve cell identity predictions, we developed a method, termed reversed hierarchical classification and repetitive elimination of parental nodes (RHC-REP) which reduces the breadth of cell types in an ensemble of classification tasks. As compared with top-down local hierarchical methods, here, predictions of terminal classes are repeatedly being evaluated to infer the cells' broader parental nodes. During this process terminal cell classes are investigated iteratively using multiple sets of refined differentiating features until confident terminal annotations are achieved.

To develop this approach, we structured a pan cancer TME cellular hierarchy where each parent node represents a group of related cells, and each terminal node represents a single-cell class of interest. Overall, we trained 24 random forest models corresponding to the total number of parent nodes (Fig. 1a). For every model, we selected DEGs that distinguish each cell type from all other terminal classes nested within the same parent. RHC-REP will then prioritise the features with the highest specificity to the interrogated cell types (Fig. 1b).

During each classification task, every cell receives a vector of prediction scores (PS) corresponding to the percentage of trees voting for each terminal class in the parent node (Fig. 1c). This cell by PS matrix is then used to calculate intermediate group scores (IGS), to subsequently link cells to their next parental node in the hierarchy (Fig. 1d, Supplementary Fig. 3). At each classification task, the distribution of IGSs obtained from all the cells interrogated in the model is used to automatically define prediction cut-offs (Supplementary Fig. 4). Each cell is then interrogated by its next associated model, defined by a more discriminative set of features and fewer potential terminal classes (Fig. 1e). Cells that do not pass the IGS thresholds are given their previous parent classification and withheld from further subclassification (Supplementary Note 1).

Given that non-malignant cells that share the cancer's tissue of origin can be found in cancer biospecimens (for example, normal alveolar cells in a lung biopsy) we embedded a cancer signature scoring and cell differentiating module in scATOMIC. Using an established transcriptional program scoring method[11,16], cancer-type-specific up and down-regulated programs[31] are evaluated in cells receiving an original annotation of a cancer type by scATOMIC (Fig. 1f).

Performance evaluation and validation across internal and external datasets

To evaluate scATOMIC's performance, we first conducted fivefold cross validation using the training reference dataset (Supplementary Data 1) while keeping equal proportions of cell types in each of the five folds. scATOMIC achieved median F1 scores ranging from 0.90 to 0.99 across all the tested cell types (Fig. 2a), implying great accuracy in classifying the breadth of cells in the settings of pan cancer TMEs. To ensure scATOMIC was not heavily impacted by batch effects, we also trained 4 scATOMIC iterations, where all cells from intact technical batches were held out from training. We found no significant difference in the performance of the models on the held out batches (Supplementary Fig. 5a, Kruskal–Wallis rank sum test: H statistic(degrees of freedom = 3) = 6.12, P value = 0.106). We further tested scATOMIC performance across different scRNA platforms using external melanoma datasets[29,32–34] and again, found no significant difference in F1 scores (Supplementary Fig. 5b, Kruskal–Wallis rank sum test: H statistic(degrees of freedom = 3) = 2.18, P value = 0.537).

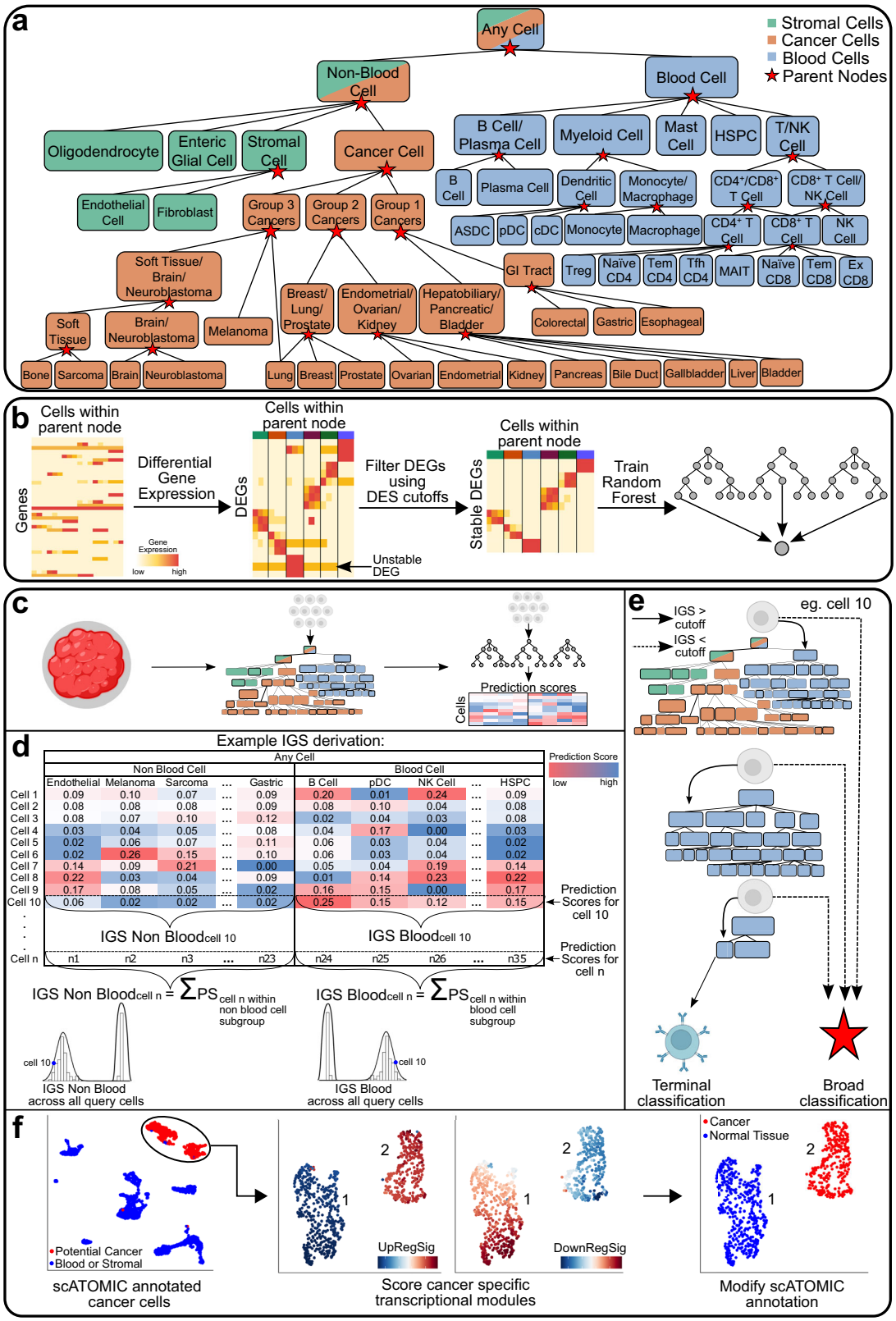
We next aimed to conduct a comprehensive external, training-independent validation of scATOMIC performance. To build a validation dataset with high-confidence cell annotations, we mined publicly available scRNA-seq data from primary tumour biopsies and blood samples. Overall, the curated set used for validation contained 228,460 cancer, 82,976 stroma and 46,090 blood cells from 225 primary biopsies spanning 13 cancer types (Supplementary Data 3). Importantly, these ground truth sets include cancer cells supported by abnormal CNV profiles, and immune cells with transcriptomic-independent identity supported by cell surface protein markers via CITE-seq. Similar to the results obtained from internal validation, in this independent validation process, scATOMIC achieved a median F1 score of 0.99 (Fig. 2b).

Overall, these results demonstrate the broad abilities of scATOMIC's core algorithm to detect cancer cells and their type, as well as predicting non-malignant cell types and subtype.

Comparison with other cell-type classification methods

We compared scATOMIC's performance to six commonly used scRNA-seq classifiers (SingleR[15], Seurat[16,35], SingleCellNet[36], scmap-cell[17], CHETAH[37], and scType[38]). These annotators encompass a variety of methods including reference correlation, label transfer using integration to a reference, random forest algorithm, flat and hierarchical models, and marker-based classification methods making their

https://doi.org/10.1038/s41467-023-37353-8



comparison with scATOMIC's underlying RHC-REP approach highly suitable. Each tool was provided with the same reference and external validation dataset as scATOMIC for training and testing (Supplementary Data 1–3). All classifiers were highly accurate in classifying non-malignant cells (median F1 > 0.85, Fig. 2c). However, for cancer cells in particular, F1-scores were significantly lower as compared with scATOMIC (two-sided Wilcoxon rank sum tests, all P values < 2×10^{-16}). The next best performing classifier following scATOMIC was SingleR[15] with median F1 scores of 0.92, 0.97 and 0.76, for blood, stromal, and cancer cells, respectively (Fig. 2c). Using scATOMIC, we obtained median F1 scores of 0.95 0.99, and 0.99 in each of these respective categories. As existing cell type classification tools were not designed to annotate malignant cells, this comparison highlights scATOMIC's ability to overcome the complexity presented in pan cancer settings to

Fig. 1 | Overview of scATOMIC training and classification. a Hierarchical structure of the pan-cancer tumour microenvironment. The cellular hierarchies in the pan-cancer tumour microenvironment are organized into a flow chart with increasing cell type resolution. Parent nodes represent broad classification branches, and terminal nodes represent specialised cell classes of interest. **b** Training of classification branches for each parent node ($n = 24$). The reference datasets are filtered based on transcriptomic-independent information to only include terminal cell types that are found within a particular parental node. Genes that significantly differentiate one cell type from all the others are gathered. Differentially expressed genes (DEGs) with greater specificity to each terminal class, determined by differential expression score (DES), are kept (Methods). A random forest classifier is trained on filtered, library size normalised count matrices to derive a model that provides prediction scores corresponding to the proportion of trees voting for each terminal class within the parental node. Colours on the top of the heatmap illustrate different cell types. **c**–**f** Classification of query datasets. **c** Gene expression count matrices from query tumour biopsies are inputted into the first scATOMIC classification branch model, outputting a cell-by-prediction scores matrix. **d** Prediction scores (PS) from all blood and non-blood cell subtypes are respectively summed to derive intermediate group score (IGS) distributions associating single cells with their appropriate parental class. **e** Cells are iteratively interrogated at their next parent nodes' corresponding models until terminal classification are obtained. Broad classifications occur if the IGS for a cell is lower than the confidence cut-off. In this example, cell 10 is subclassified until a terminal B cell designation is derived. **f** Differentiating between cancer and tissue-specific non-malignant cells through scoring of bulk RNA-seq derived differentiating gene expression programs (Methods). scATOMIC automatically annotates population 2 as cancer cells, and population 1 as non-malignant. Heatmaps and cell illustrations were created with BioRender.com.

accurately identify cancer cells while also having comparable or significantly better performance in annotating stroma and blood. Lastly, with the exception of CHETAH[37], mostly comparable time and memory usages were measured for all the methods (Supplementary Fig. 6).

Distinguishing between non-malignant, tissue-specific cells and cancer

Aneuploid CNV profiles are highly associated with the development and progression of numerous cancers by impacting gene expression levels[39]. We assessed scATOMIC's ability to distinguish malignant cells from other normal cells of the TME, sharing the same cell-of-origin, by comparing scATOMIC's final cancer predictions (Fig. 1f) to their inferred CNV-based ploidy status[19]. We observed strong agreement in cells predicted as malignant and aneuploid-inferred CNV profiles across biopsies, as well as between non-malignant detected cells and diploid inferred profiles, with a median agreement rate of 85.9% (Fig. 3). In addition, in silico serial dilution analysis of cancer cells in our external dataset collection revealed that decreasing number of malignant cells can be appropriately annotated, with scATOMIC also providing cancer type notations (Supplementary Fig. 7). Discordant cases, defined as cases with an agreement rate below the 1st quartile (Q1 ≤ 63.2%), were typically attributed to tumours with low CNV levels, intra-tumoural malignant subclones, low number of cancer cells harbouring the CNV, or low number of reference cells (Supplementary Fig. 8). In only 7% of discordant cases more cells were inferred as aneuploid than cells annotated as malignant by scATOMIC (Fig. 3b). These results suggest that cancer and related normal tissue cells are efficiently classified by their transcriptomic profiles using scRNA-seq data, independent of their ploidy status.

scATOMIC annotations increase cellular resolution in tumour biopsies

To further demonstrate the benefits of scATOMIC in annotating multicellular TMEs, we analysed several datasets, including scRNA-seq of lung cancer[29]. Original annotations for this dataset were determined by the authors using SingleR[15] with its default references in combination with cell type signatures and the use of canonical marker genes. Similar to our observations (Supplementary Fig. 1) Slyper et al.[29], noted overlapping expression programs between T cells and NK cells which makes high resolution single-cell discrimination among these cell types challenging. scATOMIC resolved NK cells and T cells, and further subclassified the latter into fine grained subtypes including T regulatory cells, naive CD4 + T cells, CD4 + T follicular helper cells, effector/memory CD4+, effector/memory CD8 + T cells, and exhausted CD8 + T cells (Fig. 4a). In addition, in this lung dataset, scATOMIC identified other distinct cell types including plasma cells, and plasmacytoid dendritic cells (pDCs) which scATOMIC separated from B cells. Unsupervised clustering and expression of cell specific markers[34] supported the existence of these separated cell identities (Supplementary Fig. 9). Of note, this biopsy included two more small clusters of "epithelial cells" (Supplementary Fig. 9a), suggesting tissue-specific cell classes that are not represented in the current scATOMIC training reference. Using scATOMIC's automatic approach to set confidence IGS cut-offs (Supplementary Fig. 4) these cells were abstained from being falsely annotated and correctly assigned with the lower-level annotation of non-blood cells. scATOMIC resolved the remaining epithelial cells into lung cancer and normal tissue cells by evaluating lung cancer associated transcriptional signatures (Fig. 1f).

Increased cellular resolution across the cell types of the TME was also observed in other recent datasets of different cancer types including bladder[4], breast[40], liver[41], ovarian[40], prostate[42], and skin cancer[33] (Fig. 4b–g). In addition, scATOMIC identified hematopoietic stem/progenitor cells (HSPCs) in glioblastoma[43] (Supplementary Fig. 10); a population which was shown to promote tumour cell proliferation[43].

Collectively, this analysis demonstrates the ability of scATOMIC's core hierarchical algorithm to resolve cell identities at high resolution, label fine grained T cell states, identify rare cell types, abstain from falsely classifying unknown cells, and determine cancer types.

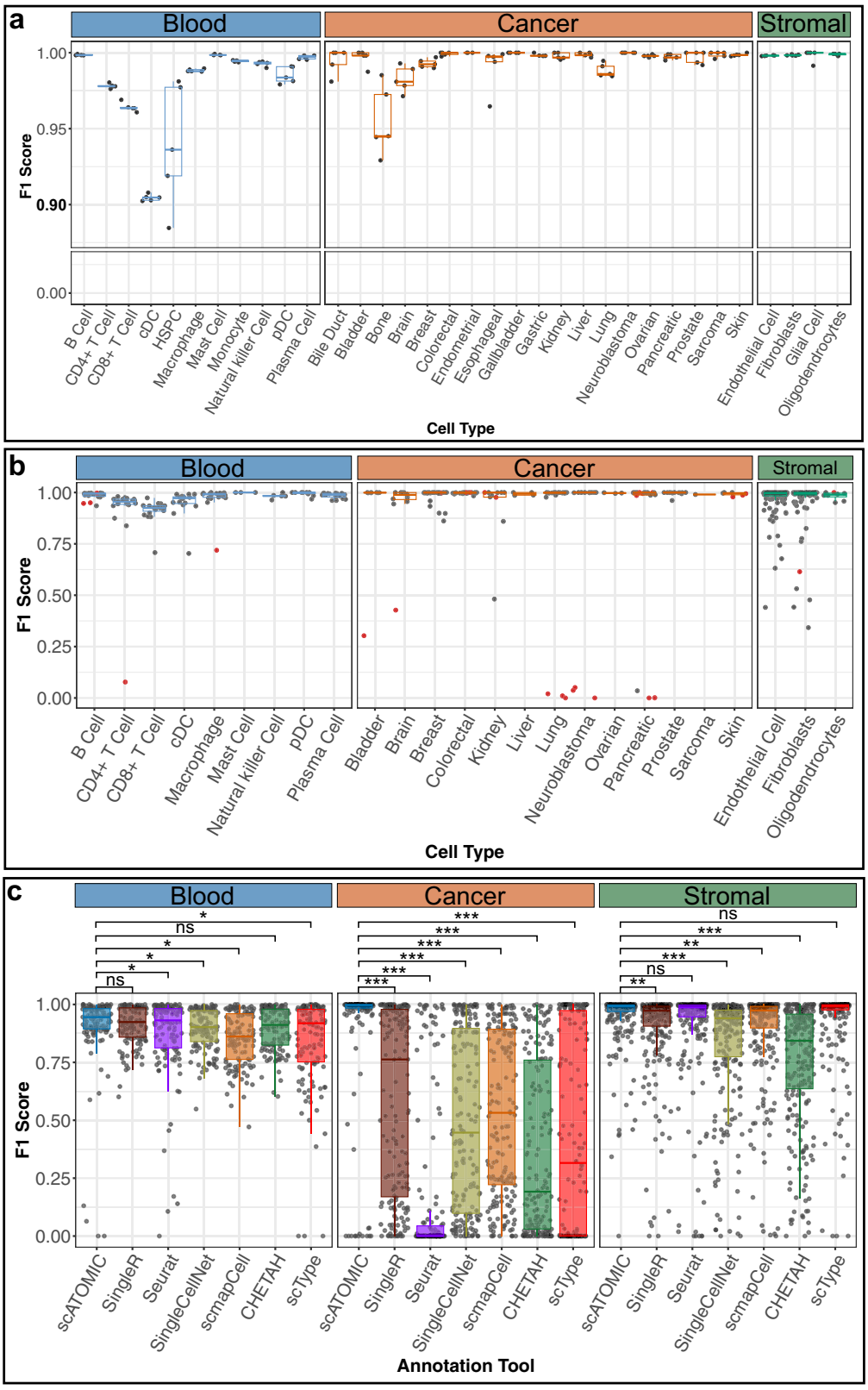
Extending the core scATOMIC hierarchy for novel applications

By leveraging RHC-REP, one can easily deploy new scRNA-seq data to train extensions at any terminal branch of the hierarchy. We thought that extending the breast cancer classification node would provide a practical example of utilising modularity (Fig. 5a). Two sizable scRNA-seq breast cancer atlases were used to train, and independently test (Supplementary Data 4, 5) a classification model that resolves breast cancer cells into the major ER + , HER2 + and triple negative breast cancer (TN) histological subtypes. We applied scATOMIC to the training-independent validation dataset containing 38 tumours spanning ER + , HER2 + , and TN breast cancer, and 2 HER2 + /ER + double positive tumours, a class not represented in the current reference of scATOMIC's breast mode due to a lack of data. scATOMIC correctly subtyped 37 of the 38 (97.4%) training-independent breast cancer biopsies, as determined by immune-staining[21,44] (Fig. 5b, Supplementary Data 5). In the two HER2 + /ER + double positive samples, scATOMIC assigned mixed annotations of HER2 + and ER + cells (Fig. 5b).

We observed different degrees of tumour cellularity, with 6 biopsies (15%) having more predicted normal breast cells than cancer cells. In another tumour reported as ER^{low} (that is, <10% ER + cancer cells by immunostaining), scATOMIC identified 8% ER + breast cancer cells (Fig. 5c, Supplementary Data 5). Of note, scATOMIC identified these ER + cells as malignant, in line with the histology report, however CNV inference predicted a diploid profile (Fig. 5d). This example highlights a distinct subpopulation of cancer cells that could have been misinterpreted as normal tissue by strictly relying on CNV inference, thus suggesting integrative approach for best results.

Overall, these data demonstrate scATOMIC's practical and modular framework to further subset primary tumour classes into their clinically relevant subtypes.

https://doi.org/10.1038/s41467-023-37353-8



scATOMIC identifies the tumour of origin across metastatic cancers

Given that existing single-cell annotation tools are not designed to provide information regarding the originating tissue of a cancer cell, we applied scATOMIC to predict the tumour origin in settings where it may be unknown. We curated a dataset of 62 metastatic biopsies from breast, kidney, lung, ovarian, and skin cancers from diverse anatomical sites (Supplementary Data 6). In 52 of the 62 samples (83.9%), the primary tissue of origin was correctly predicted by scATOMIC (Fig. 6), demonstrating its robustness at distant sites, in cells that may have undergone transcriptional changes associated with metastasis. In 1 kidney and 2 lung samples (additional 4.9%) scATOMIC abstained from giving a terminal classification yet focused the prediction on the correct intermediate class. In 2 lower throughput melanoma scRNA-seq,

Fig. 2 | scATOMIC performs accurately in internal and external validation experiments. a *k*-fold cross validation. The reference dataset was randomly split into 5 sub-samples containing equal numbers of each cell type. F1 scores are shown for each cell type in each of the 5 replicates (jitter points). Each fold contained overall ~61,100 cells. Boxplot colours represent the major cell type classes. **b** External validation in datasets not used for training. scATOMIC was validated on CITE-seq datasets of tumour derived blood cells, datasets of aneuploid cancer cells and stromal cells from primary tumour biopsies. F1 scores are shown for each cell type within individual samples (jitter points). The plot represents $n = 357{,}526$ cells from 225 samples. Red dots indicate low-confidence cell type classifications (Methods). Boxplot colours represent the major cell type classes. **c** scATOMIC outperforms other existing automatic cell type annotators, particularly when applied to identify cancer cells and determine their type. Six existing classifiers were provided the same training/reference and training-independent validation datasets as scATOMIC. Combined F1 scores for each of the three major cell class, blood, cancer, and stroma are shown (jitter points). The plot represents $n = 337{,}790$ cells from 221 samples that were given a classification output by all tools. (two-sided Wilcoxon rank sum test comparing scATOMIC to each tool $^{*}P < 0.05$, $^{**}P < 1.1 \times 10^{-6}$, $^{***}P < 2 \times 10^{-16}$, are shown). Boxplot colours represent the different tools. For all plots, boxes and whiskers represent the lower fence, first quartile (Q1), median (Q2), third quartile (Q3), and upper fence. Source data are provided as a Source Data file.

only 5 and 6 cancer cells were reported[11] yet scATOMIC found none. We considered these as false predictions. In 4 of the 5 remaining samples that received incorrect terminal classifications, the predicted cancer type and the reported primary were related cancers falling under the same immediate parent node. For example, a mixed serous/clear-cell ovarian carcinoma was predicted to be endometrial cancer, with relatively low classification scores (Supplementary Fig. 11). Overall, these results show that accurate detection of metastatic cancers' tissue of origin using single-cell transcriptomics is feasible and that scATOMIC can aid in identifying cancers' primary sites across a variety of solid human tumours.

Discussion

The rate of scRNA-seq publications reporting major scientific insights concerning the function of various immune and stromal cells in cancer has increased steadily over the years[40,45,46]. However, the lack of automated methods that can also standardise the identification of single malignant cells is becoming a major obstacle to accurately study tumour-microenvironment interactions across various cancer types.

We developed scATOMIC to effectively annotate the TME in pan-cancer settings. scATOMIC overcomes several classification challenges, including high inter-patient heterogeneity and highly overlapping expression profiling among specificized immune cells. By using stably expressed transcripts as features, structured classification, and models trained using reliable and large datasets, scATOMIC has proven to accurately identify cancer cells and their origin. Moreover, scATOMIC is comparable to or outperforms other existing automatic cell type annotators when classifying blood and stromal cells using our training reference. In samples with genome instability and an appropriate reference of normal cells, we found high agreement between scATOMIC and CNV inference to pinpoint malignant cells in scRNA-seq data. However, in samples with cancer sub-clones defined by variable CNV burdens, cancer cells with near diploid genomic profiles, or few normal cells to serve as controls, CNV inference may fall short. Since information concerning CNV may be useful for cancer prognostication, and a degree of discordancy still exists, using scATOMIC in conjunction with CNV inference to annotate cancer cells and their type is recommended.

We designed the core, ploidy-neutral scATOMIC algorithm to accurately identify cancer and normal tissue cells across 19 common cancer types, including key rare populations such as plasmacytoid dendritic cell and hematopoietic stem and progenitor cells that were reported in cancer tissues and are associated with immunosuppressive phenotypes[29,43]. To ensure that scATOMIC remains powerful, we designed it in a way that new data can be easily interrogated to extend the core hierarchy by adding new terminal cell classes. We demonstrated this modularity by further classifying breast cancers into their clinically relevant molecular subtypes achieving high agreement between transcriptomics and immunostaining. With the progressive accumulation of high quality publicly available scRNA-seq data, future extensions of the core hierarchy to further subclassify the other 18 cancer types and the various core, non-malignant cell types to their more resolved classes or states will become simple.

As molecular classification of cancers by tissue-of-origin is fundamental to diagnostic pathology we demonstrated scATOMIC's ability and high accuracy in predicting the primary origin of metastatic tumours. Additional work is required to evaluate the limits of single-cell transcriptomics to predict cancer origin, specifically in cancers of unknown primary and other contexts where distinguishing primary from metastatic tumours is not trivial, such as in the case of mucinous ovarian carcinoma[47].

In summary, we have described, benchmarked, and validated a highly accurate single-cell annotation tool across TMEs of common, deadly cancer types. The RHC-REP classification approach underlying scATOMIC used here to tackle the complexity associated with multicellular TMEs might be of interest in areas outside the cancer field entailing multi-class structured systems. We highlight the benefits that scATOMIC holds in cancer settings compared to other tools, providing a method to standardise single-cell cancer transcriptomic studies. We expect that scATOMIC's abilities to accurately identify TME resident cells with high resolution, separate between cancer and normal tissue cells, and determine tumours' origin will enrich and expedite broad cancer studies seeking to refine prognostication or cell–cell communication from single-cell transcriptomes.

Methods

Defining the pan-cancer tumour microenvironment cell type hierarchy

We defined a structured hierarchy where cell types with transcriptomic similarities are grouped into nodes (Fig. 1a). We first grouped blood cells based on existing relationships that correspond to the hematopoietic hierarchy[48], and kept stromal cells together. For cancer cells, we derived putative groups where transcriptomic similarities are expected based on the cancers' shared organ system, histological subtype, hormonal tissue, or germ layer. These included carcinomas of the digestive system (group 1: colorectal, gastric, esophageal, liver, gallbladder, bile duct, and pancreas), carcinomas not of the digestive system (group 2: lung, breast, prostate, endometrial, and ovarian), and non-carcinomas including soft tissue, neuroendocrine, and nervous system cancers (group 3 cancers: bone, sarcoma, brain, melanoma, and neuroblastoma). We then used a subset of our data, corresponding to one patient sample from each cancer type to evaluate random forest models. Evaluating the proportion of trees voting for each cancer type helped refine the groups and provided guidelines concerning what the subsequent nodes might be.

For example, for the less trivial grouping of kidney cancer, a classification model including all cancers assigned 42.3% of tree votes for kidney, while the majority of the remaining trees voted for various group 2 cancers thus, suggesting linking kidney cancer with group 2. In the other case of lung cancer, we decided to include it in both groups 2 and 3, as separate cancers from both these groups obtained a high number of trees voting for lung. In this case, no clear distinction was observed. In a different case concerning CD8 + T cells, we included CD8 + T cells in both child nodes of the parent node T/NK as different CD8 + T cell populations show more transcriptomic similarities to NK

https://doi.org/10.1038/s41467-023-37353-8

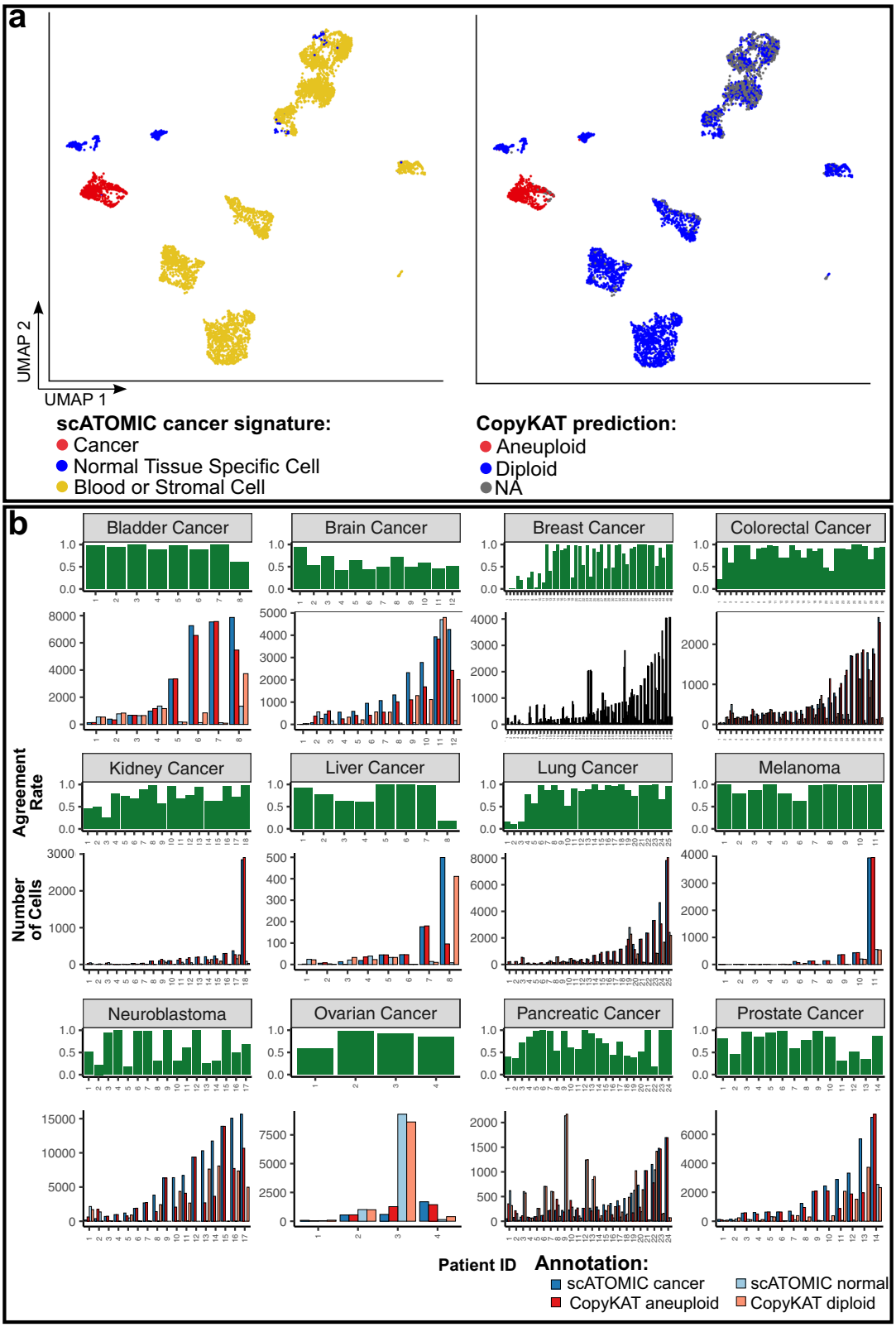


Fig. 3 | scATOMIC effectively distinguishes between malignant cells and normal tissue specific cells. a scATOMIC predictions and inferred ploidy in breast cancer patient CID4066[21]. Cells are coloured by scATOMIC predictions and copy number variation (CNV)-based inferred ploidy. scATOMIC-predicted malignant cells are inferred as aneuploid cells while normal tissue cells are inferred as diploid. **b** Comparison of scATOMIC cancer predictions and inferred ploidy statues across the training-independent, external validation datasets. Blue bars represent the number of cells predicted as malignant (solid blue) and non-malignant (transparent blue) by scATOMIC. Red bars represent the number of cells inferred as aneuploid (solid red) and diploid (transparent red). Green bars represent agreement rate in each biopsy. Rates do not include cells without a confident ploidy status (that is received an "NA" annotation by CopyKAT). Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-023-37353-8

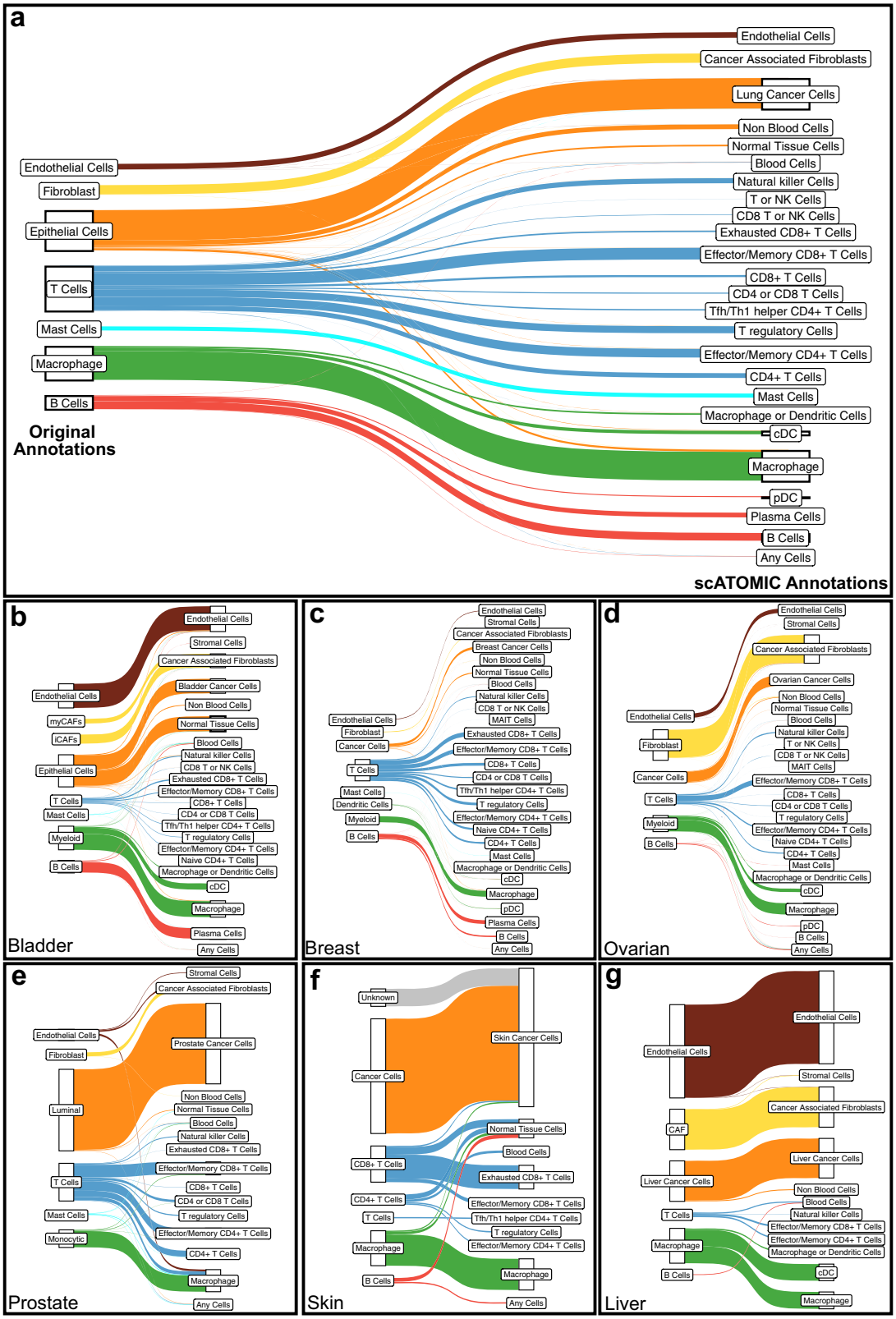


Fig. 4 | scATOMIC provides greater cellular resolution than original annotations across tumour datasets. a Sankey plot comparing original cell type annotations to higher resolution scATOMIC annotations in a recent lung cancer biopsy dataset[29]. scATOMIC identifies lung cancer as the tissue of origin and distinguishes these cells from normal lung tissue cells. scATOMIC identifies subtypes of blood cells. **b**–**g** scATOMIC identifies the tumour origin of common cancers and deliver relatively higher resolution in other cell types[4,33,40–42]. Colours represent the original reported annotations associated with each dataset. The height of each block represents the relative number of cells that received a respective annotation. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-023-37353-8

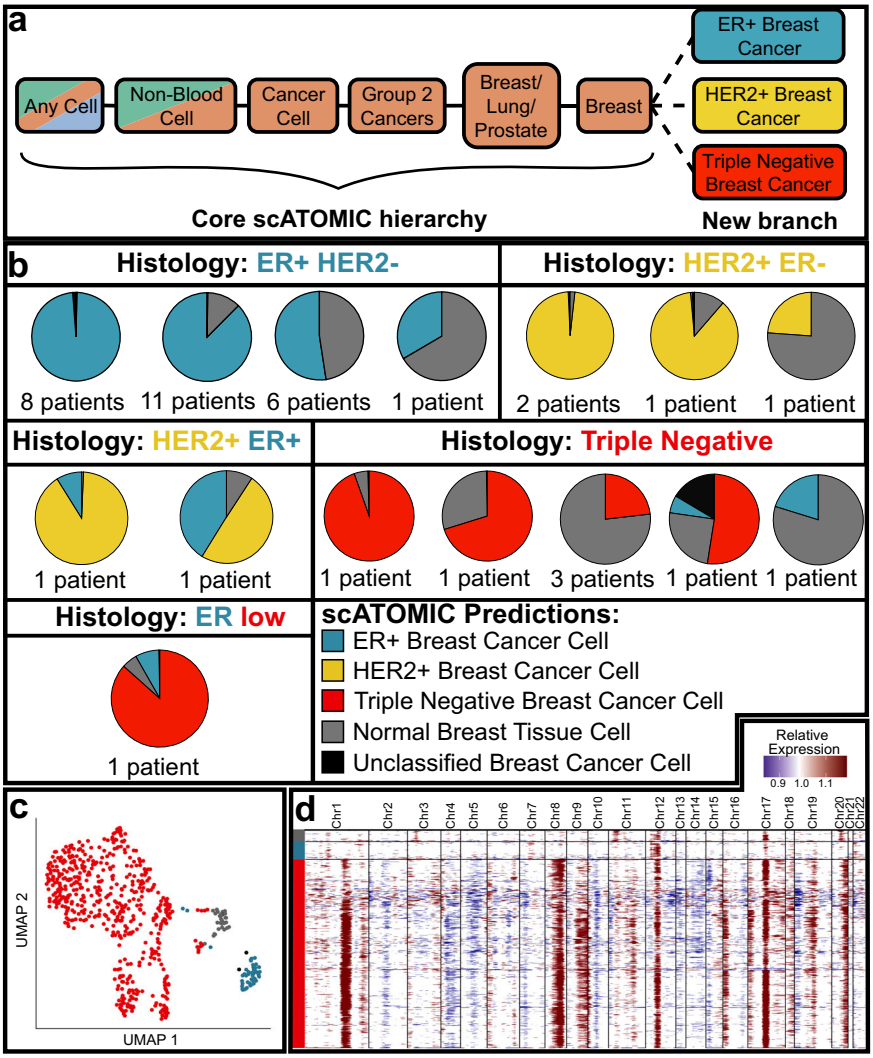


Fig. 5 | Extending the core scATOMIC model to further classify breast cancer subtypes. a The terminal breast cancer cell node from the core hierarchy of scATOMIC is extended to subclassify breast cancers into their major ER+, HER2+, and triple negative histological subtypes. **b** Validation of scATOMIC predictions in an external cohort[44]. Pie charts reflecting intra-tumoural breast subtype heterogeneity according to scATOMIC classification are shown for each reported histological subtype. Patient specimens with similar distributions of cell annotations are illustrated together in a single pie chart. **c** Breast cells from an ER-low tumour (Patient: ER-AH0319) are visualised on UMAP and coloured by scATOMIC predictions. ER + breast cancer cells represent a sub-clonal cancer cell population. **d** Inferred copy number variation (CNV) profiles of cells from ER-low tumour. Red represents inferred gains, while blue represents inferred losses of genomic regions. The *y* axis is coloured according to scATOMIC prediction. Colours representing scATOMIC predictions apply to all the panels in the figure. Source data are provided as a Source Data file.

(i.e. cytotoxic T cells) while others resemble CD4 + T cells more (Supplementary Fig. 1). We found that this structure yields stronger performance as compared to only having CD8+ cells in one of the branches or generating a single model to differentiate between CD4+, CD8+ and NK all at once. Overall, given that a small random data subset was utilised to infer transcriptomic similarities among classes of cells, it is possible that other hierarchical structures might also be appropriate.

Feature selection

For every classification branch within the core scATOMIC hierarchy we selected features as a training input of a random forest model. Raw gene by cell count matrices derived from scRNA-seq analysis were gathered from multiple sources (Supplementary Data 1) and were organized into 24 parent groups (Supplementary Fig. 3). To merge matrices into a particular parent dataset, we removed genes that are not represented in all of the data sources. In each parent dataset we removed cells with <500 expressed genes (as defined by non-zero counts) or with more than 25% of their reads being mapped to mitochondrial genes.

To find DEGs between each terminal cell type and all other terminal cell types present in the same parent node we used the 'Seurat' R package v4.0.1[16]. Raw gene by cell count matrices were normalised and variance stabilised using the SCTransform function to remove technical variability. Principle components were found using the RunPCA function, on the "SCT" assay. Louvain clustering was performed by first applying the FindNeighbors function on the top 50 PCs, followed by the FindClusters function with a resolution of 2. The identity of the resulting cell clusters was determined by the transcriptomic-independent ground truth associated with the training datasets. For each model, DEGs were found using the FindMarkers function (two-sided Wilcoxon rank sum test) with ident.1 set to include all clusters containing a particular terminal cell type and ident.2 being all other cells in the parent node. The function returned a list of DEGs per class that passed default Seurat filtering settings: a $\log_2$ fold-change of at least 0.25, and at least 10% of the cells in ident.1 or ident.2

https://doi.org/10.1038/s41467-023-37353-8

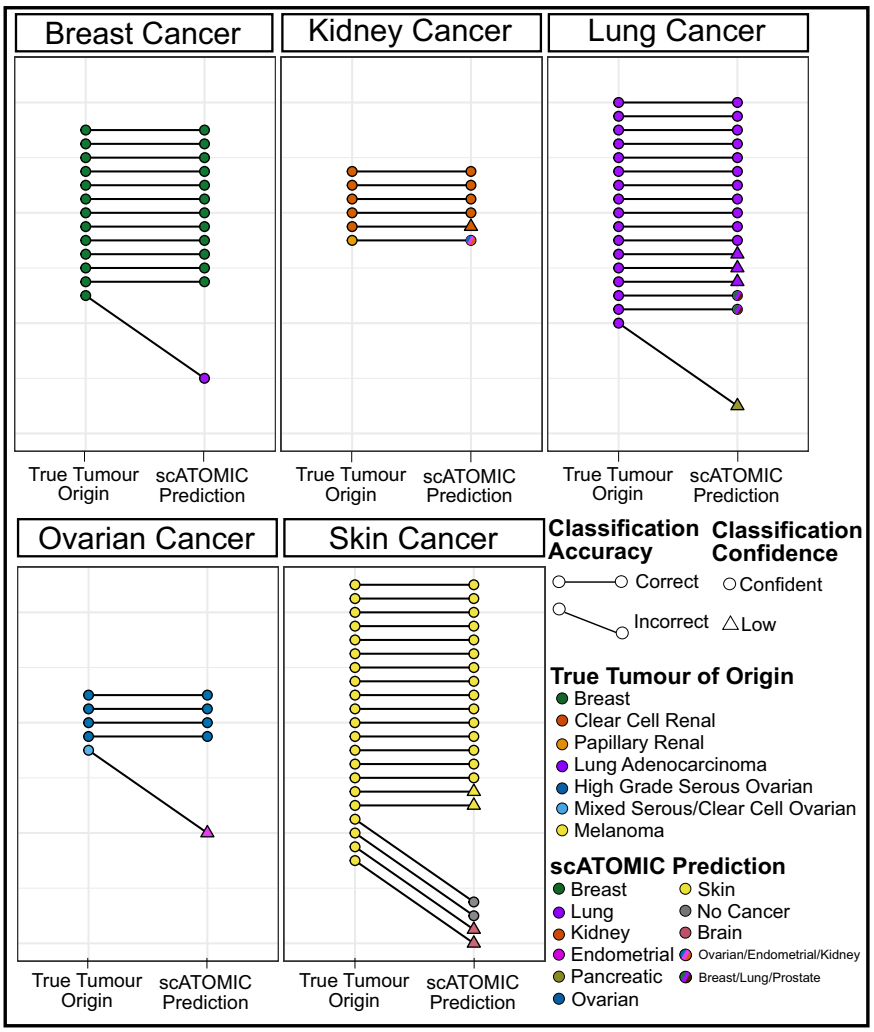


Fig. 6 | scATOMIC accurately identifies the tissue of origin in metastatic tumour biopsies. scATOMIC was applied to 62 metastatic tumours from breast, kidney, lung, ovarian and skin. Metastatic sites included the brain, lungs, GI tract, liver, adrenal glands, lymph nodes, abdomen, and peritoneal cavity. Each pair of dots represents the true tumour origin and the predicted origin. Horizontal connected lines represent correct predictions, while diagonal lines represent incorrect predictions. True tumour origins are coloured by the reported cancer subtype. Circular points represent confident annotations, while triangular points represent low-confidence annotations (Methods). Multi-coloured points represent tumours that received an intermediate scATOMIC annotation. Source data are provided as a Source Data file.

expressing the respective gene. We defined a differential expression score (DES) as the difference of the fraction of cells expressing a non-zero value for a respective DEG in ident.1 and ident.2 (DES = pct_expr_ident.1 − pct_expr_ident.2) to capture DEGs more specifically expressed in any particular cell type. For each terminal cell type, we kept genes with a DES greater than the mean DES of all DEGs for that cell type. We removed all ribosomal genes. We also removed DEGs that had a pct_expr_ident.2 >40% to ensure high performance when interrogating datasets with large technical variation. For the same reason, we set a minimum and maximum number of DEGs for each cell type at 50 and 200. Specifically, a minimum number of features was set to mitigate potential issues in classifying cells with high levels of technical dropout. In the case where there are fewer than 50 DEGs with DES higher than the mean, we kept the top 50 DEGs ranked by DES. Features that were used for each cell class at each classification layer are provided in Supplementary Data 7.

Random forest modeling

For each classification branch within the core scATOMIC hierarchy, we trained a random forest model on cells within a respective parental node and features selected as described above. To minimise bias associated with imbalanced classification towards majority classes[49], we randomly sampled an equal number of cells from each terminal class, with replacement. Library size of each single cell was normalised by using the library.size.normalise function from the 'Rmagic' v2.0.3 package[50]. Prior to training each model, normalised counts were filtered to include selected features and cells within the corresponding parental node. Read count values were transformed to a fraction of the total filtered counts. A random forest classifier was trained on the transformed matrix using the 'randomForest' R package v4.6–14 with 500 trees and default parameters[51]. Each random forest was trained to classify the terminal nodes present within the corresponding parental node. The specific cell type organization of the 24 classifiers is detailed in Supplementary Fig. 3.

Applying scATOMIC to query datasets

Before using scATOMIC on query datasets, the interrogated data was processed as follows. Raw gene by cell count matrices were filtered to remove cells with non-zero counts for <500 genes or with more than 25% of their reads being mapped to mitochondrial genes. We imputed missing values in DEG using the *magic* function from the 'Rmagic' package[50], where all the cells within the query dataset act as a

reference. In datasets where there were no reported values across all the samples for specific selected features, we assigned a value of zero before imputation. Following each classification task, every cell received a vector of prediction scores corresponding to the percentage of trees voting for each terminal class in the random forest model. The values in each vector within the next immediate parent node were then summed to generate IGSs. For example, when the first model interrogating all the terminal classes in the hierarchy ran (i.e. the parent model "Any Cell"), the output for each cell was composed of two intermediate group scores. The first corresponded to the sum of trees voting to all the terminal cell classes belonging to the "Blood Cell" parent node and the second IGS for the "Non-Blood Cells" parent node (Fig. 1d). Data from all the cells in the interrogated sample was used to derive IGS parent distributions. Cells which received an IGS greater than the defined parent threshold continued down the classification hierarchy until they were terminally classified (Fig. 1e). At any stage, if a cell received an IGS lower than the calculated threshold it was annotated based on the previous parental node (a less specific classification). An IGS threshold for a classification to be deemed confident was automatically determined (Supplementary Fig. 4). Using the ajus function from the 'agrmt' package v1.42.4[52], IGS distributions for each IGS calculated among all cells within a parental node are classified as either unimodal or bimodal. Unimodal distributions suggest a layer includes one subtype, while bimodal distributions indicate there is likely more than one subtype. For unimodal distributions we set the IGS threshold to be the mean IGS (μ) − 3 standard deviations (σ). For bimodal distributions, using the em function from the 'Cutoff' R package[53] v0.1.0, we fit a mixture model for the distributions and predicted estimates of mean and standard deviations for both distributions using the expectation maximation algorithm. We selected a conservative approach when a mixed cell type population exists in a layer by setting the IGS threshold in bimodal distributions to be the mean of the higher score modality (μ_2) − 2 standard deviations (σ_2). We set the maximum IGS threshold to be 0.7, as in some distributions, such as when a single pure population remained for classification, unreasonably high thresholds may be obtained.

A schematic and description of the main functions is detailed in Supplementary Note 2.

Flagging cells with lower confidence annotations

To provide a way for scATOMIC users to evaluate the confidence of their cell annotation output, we devised a secondary post-classification flag to warn about low-confidence annotations. To define low-confidence annotations, we used the results obtained by external validation (Fig. 2b). For every model throughout the hierarchy, we determined the median IGSs (or PS for terminal nodes) across samples for correctly annotated cells (X) and incorrectly annotated cells (Y). Correct versus incorrect status was defined by the terminal annotation of single cells. For example, in terminally annotated breast cancer cells, median $IGS_{non\text{-}blood}$, $IGS_{non\text{-}stromal}$, $IGS_{group2\text{-}cancer}$, $IGS_{breast/lung/prostate}$, PS_{breast} for all the correctly and incorrectly annotated cells in the validation data were recorded. We derived confidence thresholds based on the overlap between the distributions of X and Y using their quartiles (Q). When there was low overlap (defined as minimum X > Q3 Y), the threshold was set to the minimum X. When there was intermediate overlap (defined as minimum X < Q3 Y, yet Q1 X > Q3 Y), the threshold was set to Q3 Y. In all remaining cases where there was high overlap (defined as Q1 X < Q3 Y), segregation could not be made and the classifications of query cells in such cases were deemed confident. For example, all CD4 + T cells that are misclassified as CD8 + T cells will still obtain comparable IGS_{blood} with respect to correctly classified CD8 + T cells, both being subtypes of blood cells. If at any model throughout the hierarchy a low-confident IGS is observed, a flag is applied. In addition, we assigned a sample-level confidence metric derived from the proportion of cells receiving a confident annotation based on this flag. To maximise identification of potentially poor-quality samples, in any new interrogated sample, if <75% of cells receive confident annotations, a warning will be issued (Supplementary Fig. 12).

Scoring cancer signatures to refine cancer cell predictions and identity of normal tissue-specific cells

To identify potential normal tissue specific cells that are not defined in the core scATOMIC TME hierarchy, we employed a post classification method for scoring cancer specific modules in each cell. Lists of genes differentially expressed between different cancer types and their matched normal tissues were obtained from OncoDB[31]. We selected DEGs from OncoDB[31] with a reported log2 fold change >1 or <−1 and an adjusted P value <0.01. Since OncoDB[31] is based on bulk RNA-seq, we further filtered the DEG list to only include those with reported expression values in the query scRNA-seq dataset. Upregulated gene programs and downregulated gene programs[31] were scored using the AddModuleScore function from Seurat[11,16] in each cell predicted as cancer by random forests. Ward.D2 hierarchical clustering was then performed on a Euclidean distance matrix of each cell's upregulated and downregulated cancer programs. Two groups were derived using the cutree function. At this stage, scATOMIC evaluates the percentage of normal cells in each group corresponding to those cells annotated as either blood, stroma or cancer cells with lower upregulated program scores compared to downregulated program scores. As the AddModuleScore function uses average expression of control feature sets across all cells in the dataset, the calculated scores are affected by the proportion of normal cells present. Thus, we filtered out all confident normal cells in the cluster with a greater percentage of normal cells and repeated the AddModuleScore pipeline to identify additional normal cells that were overlooked in the first iteration. We repeated scoring of cancer programs, hierarchical clustering, calculating the percentage of normal cells and filtering until both clusters contained no more than 20% normal cells. Cells that were initially classified as cancer which were scored as normal cells were given a normal tissue cell label.

Benchmarking scATOMIC

We validated scATOMIC's performance using internal cross validation and external validation datasets. Internal validation was performed by splitting the training dataset (Supplementary Data 1) into five subsets containing equal proportions of each cell type. For each iteration we used 4 subsets as scATOMIC training dataset and applied scATOMIC to the held out independent test subset. F1 scores were calculated for each terminal cell type at each iteration (Supplementary Data 1). AXL^+ dendritic cells (ASDC) represent a rare, recently discovered, transitional state between cDCs and pDCs[54]. Due to their small numbers in the training data, these cells were omitted from the internal validation procedure.

For external validation of scATOMIC we used 424,534 cancer, blood and stromal cells gathered from various sources (Supplementary Data 3). For ground truth, we relied on the authors' annotations of stromal cells. To improve reliability, cells that were annotated as cancer by the authors were subjected to additional validation by CNV inference using CopyKAT[19]. For blood cells, we used CITE-seq data with protein surface markers supporting the author derived cell type annotations. We excluded cell types from individual samples if their number was <30. Per sample F1 scores were calculated for each terminal cell type. During both scATOMIC's internal and external validation processes (Fig. 2a, b), we considered correct intermediate classifications as true positives.

Comparison between scATOMIC and other tools

We compared scATOMIC's performance to existing scRNA-seq classifiers. In this comparison only, we bypassed the use of IGS cut-offs in

scATOMIC to enable comparison to other tools that cannot output intermediate cell classes. Of note, a separate analysis evaluating scATOMIC's performance under this forced mode against its default (unforced mode) favoured the latter by indicating that the use of intermediate annotations more frequently avoids derivation of incorrect terminal annotations than restricting the output of correct terminal classes. The same training and validation datasets provided to scATOMIC were also used to provide a comparison of scATOMIC with the performance of all the other tested tools. A SingleCellNet[36] random forest model was trained on a balanced reference of 2500 random samples of each cell type, using default parameters. SingleCellNet expects a class-balanced matrix as input. We provided the same class-balanced matrix used in scATOMIC's first classification model, representing all the cell classes. CHETAH[37] is using an alternative hierarchical classification approach. We used CHETAH with its default settings with the exception of the 'thresh' parameter that was set to zero to enforce terminal annotations, similar to scATOMIC. Scmap-cell[17] classification was performed using default parameters. To enable SingleR processing of the large reference dataset, we applied its pseudobulk implementation by setting 'aggr.ref' to TRUE[15]. For the comparison with Seurat[16,35], we partitioned the pan-cancer TME training reference according to batch and applied the reciprocal PCA workflow. We transferred labels to query samples using the TransferData function. For the comparison with scType[38], we provided scType with the list of features corresponding to each terminal class in scATOMIC's reference which were derived in the first classification node (Supplementary Data 7). Otherwise, default parameters were used. To compare scATOMIC's final cancer cell prediction with the CNV inference approach of detecting malignant cells we used CopyKAT[19] with its default settings (Fig. 3). Both aneuploid and diploid cells from the external validation biopsies were included in this analysis. Agreement rate was defined as the simple percentage agreement using the agree function from the irr[55] R package. Cells which received an NA ploidy annotation were omitted from the calculation.

Analysis of run time and memory usage

We applied each tool without parallel processing, as only some tools (scATOMIC, Seurat, SingleR) provide that functionality. We considered the time and memory usage for query datasets following model training as some methods bypass this step and rely on a given cell marker list (scType). Using the peakRAM[56] R package we monitored the time for classification and maximum RAM used. We compared the time and RAM required for the classification of cell types prior to cancer signature scoring by scATOMIC to SingleR, Seurat, scmap-cell, SingleCellNet, CHETAH, and scType final classifications. As appropriate, we compared the time and memory for the cancer signature scoring step that differentiates cancer from normal tissue cells to CopyKAT.

In silico dilution assay

For each primary tumour sample in the external validation, we ran scATOMIC and CopyKAT on all non-malignant cells while iteratively reducing the number of malignant cells. Specifically, we sampled 10–100 malignant cells in increments of 10. Only cells that both scATOMIC annotated as cancer cells and CopyKAT inferred as aneuploid were sampled.

Breast cancer subclassification

We extended the core scATOMIC model to subclassify breast cancer cells into their histological subtypes. The model extending breast cancer into its molecular subtypes was trained using a combination of two datasets. The first included the breast cancer cell lines data within the core training dataset (Supplementary Data 1, 2), where the assigned molecular subtypes were based on annotations found in the Cancer Cell Line Encyclopedia DepMap portal[57] and the second, primary tumour data from Wu et al.[21] with immunohistochemistry information (Supplementary Data 4). We used the clinical molecular subtypes of HER2 + , ER + and triple negative as classes. There was not sufficient data available to train and test HER2 + /ER + and HER2-/ER + as separated classes. We evaluated this extension's performance on an external set of 40 tumours from Pal et al.[44] (Supplementary Data 5). One tumour containing fewer than 100 cancer cells was omitted from this analysis. The ground truth of ER and HER2 status in the testing set was received by correspondence with the authors.

Visualising inferred CNVs

To visualise the inferred CNV profile in the ER-low tumour we used the 'inferCNV'[58] package with the cutoff variable set to 0.1 and all other variables set to default. We defined normal reference cells as cells annotated by scATOMIC as blood or stromal cells.

Applying scATOMIC to metastatic tumours

scATOMIC was applied to predict the tumour of origin in 62 metastatic tumours. Individual samples used are described in detail in Supplementary Data 6. We defined the scATOMIC tumour of origin prediction by taking the cancer type called in the majority of cancer cells in the sample.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the scRNA-seq data used in this work are publicly available. Datasets retrieved from the Gene Expression Omnibus can be downloaded using the following accession numbers: GSE164378[16], GSE118257[24], GSE114374[25], GSE135893[27], GSE140819[29], GSE148673[19], GSE176078[21], GSE132465[22], GSE125449[41], GSE131907[23], GSE115978[33], GSE137804[59], GSE141445[42], GSE154826[60], GSE123139[34], GSE161529[44]. Datasets retrieved from the Broad Institute Single Cell Portal are available with the following accession numbers: SCP542[20], SCP1288[61], SCP1415[32]. The remaining datasets were downloaded directly from links provided in their corresponding publications: Madissoon et al.[26] [https://data.humancellatlas.org/explore/projects/c4077b3c-5c98-4d26-a614-246d12c2e5d7], Chen et al.[4] [https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-020-18916-5/MediaObjects/41467_2020_18916_MOESM2_ESM.zip], Couturier et al.[62] [https://datahub-262-c54.p.genap.ca/GBM_paper_data/GBM_cellranger_matrix.tar.gz], Qian et al.[40] [https://lambrechtslab.sites.vib.be/en/pan-cancer-blueprint-tumour-microenvironment-0], Young et al.[63] [https://www.science.org/doi/suppl/10.1126/science.aat1699/suppl_file/aat1699_datas1.gz.zip], Peng et al.[64] [https://zenodo.org/record/3969339], Zheng et al.[45] [https://zenodo.org/record/5461803]. Additional descriptions of these datasets are provided in Supplementary Data 8. All re-processed data used for training and validation have been deposited in Zenodo and are available through the following link: https://doi.org/10.5281/zenodo.7419236[65]. Bulk RNA sequencing cancer specific signatures were obtained from OncoDB[31] [https://oncodb.org/data_download.html]. Subtypes of cancer cell lines were derived from the Cancer Cell Line Encyclopedia in DepMap[57] [https://depmap.org/portal/ccle/]. Source data are provided with this paper.

Code availability

The scATOMIC R package, associated code and user manual are available at the abelson-lab/scATOMIC GitHub repository: https://github.com/abelson-lab/scATOMIC[66]. Additional scripts to reproduce the figures in the manuscript are deposited in Zenodo and are available through the following link: https://doi.org/10.5281/zenodo.7419236[65].

References

1. Karaayvaz, M. et al. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat. Commun.* **9**, 1–10 (2018).
2. Maynard, A. et al. Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing. *Cell* **182**, 1232–1251.e22 (2020).
3. Sade-Feldman, M. et al. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* **175**, 998–1013.e20 (2018).
4. Chen, Z. et al. Single-cell RNA sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma. *Nat. Commun.* **11**, 1–12 (2020).
5. Valdes-Mora, F. et al. Single-cell transcriptomics in cancer immunobiology: The future of precision oncology. *Front. Immunol.* **9**, 2582 (2018).
6. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol*. https://doi.org/10.15252/msb.20188746 (2019).
7. Clarke, Z. A. et al. Tutorial: guidelines for annotating single-cell transcriptomic maps using automated and manual methods. *Nat. Protoc.* **16**, 2749–2764 (2021).
8. Xie, B., Jiang, Q., Mora, A. & Li, X. Automatic cell type identification methods for single-cell RNA sequencing. *Comput. Struct. Biotechnol. J.* **19**, 5874–5887 (2021).
9. Zappia, L. & Theis, F. J. Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape. *Genome Biol.* **22**, 301 (2021).
10. Fan, J., Slowikowski, K. & Zhang, F. Single-cell transcriptomics in cancer: computational challenges and opportunities. *Exp. Mol. Med.* **52**, 1452–1465 (2020).
11. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
12. Puram, S. V. et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* **171**, 1611–1624.e24 (2017).
13. Vázquez-García, I. et al. Ovarian cancer mutational processes drive site-specific immune evasion. *Nature* **612**, 1–9 (2022).
14. Domínguez Conde, C. et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* **376**, eabl5197 (2022).
15. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
16. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
17. Kiselev, V. Y., Yiu, A. & Hemberg, M. Scmap: Projection of single-cell RNA-seq data across data sets. *Nat. Methods* https://doi.org/10.1038/nmeth.4644 (2018).
18. Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
19. Gao, R. et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* 1–10 https://doi.org/10.1038/s41587-020-00795-2 (2021).
20. Kinker, G. S. et al. Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* https://doi.org/10.1038/s41588-020-00726-6 (2020).
21. Wu, S. Z. et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* **53**, 1334–1347 (2021).
22. Lee, H. O. et al. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat. Genet.* **52**, 594–603 (2020).
23. Kim, N. et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* **11**, 1–15 (2020).
24. Jäkel, S. et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).
25. Kinchen, J. et al. Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* **175**, 372–386.e17 (2018).
26. Madissoon, E. et al. scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* **21**, 1 (2019).
27. Habermann, A. C. et al. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci. Adv.* **6**, eaba1972 (2020).
28. Regev, A. et al. The human cell atlas. *Elife* https://doi.org/10.7554/eLife.27041 (2017).
29. Slyper, M. et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med.* **26**, 792–802 (2020).
30. Kiselev, V. Y., Andrews, T. S. & Hemberg, M. Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat. Rev. Genet. 2018 205* **20**, 273–282 (2019).
31. Tang, G., Cho, M. & Wang, X. OncoDB: an interactive online database for analysis of gene expression and viral infection in cancer. *Nucleic Acids Res.* **50**, D1334–D1339 (2022).
32. Wu, S. Z. et al. Cryopreservation of human cancers conserves tumour heterogeneity for single-cell multi-omics analysis. *Genome Med.* **13**, 1–17 (2021).
33. Jerby-Arnon, L. et al. A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* **175**, 984–997.e24 (2018).
34. Li, H. et al. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* **176**, 775–789.e18 (2019).
35. Stuart, T. et al. Comprehensive Integration of Single-Cell Data. *Cell* https://doi.org/10.1016/j.cell.2019.05.031 (2019).
36. Tan, Y. & Cahan, P. SingleCellNet: A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. *Cell Syst.* **9**, 207–213.e2 (2019).
37. de Kanter, J. K., Lijnzaad, P., Candelli, T., Margaritis, T. & Holstege, F. C. P. CHETAH: a selective, hierarchical cell type identification method for single-cell RNA sequencing. *Nucleic Acids Res.* **47**, e95 (2019).
38. Ianevski, A., Giri, A. K. & Aittokallio, T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. *Nat. Commun.* **13**, 1–10 (2022).
39. Shao, X. et al. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. *BMC Med. Genet.* **20**, 1–14 (2019).
40. Qian, J. et al. A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. *Cell Res.* **30**, 745–762 (2020).
41. Ma, L. et al. Tumor Cell Biodiversity Drives Microenvironmental Reprogramming in Liver Cancer. *Cancer Cell* **36**, 418–430.e6 (2019).
42. Chen, S. et al. Single-cell analysis reveals transcriptomic remodellings in distinct cell types that contribute to human prostate cancer progression. *Nat. Cell Biol.* **23**, 87–98 (2021).
43. Lu, I. N. et al. Tumor-associated hematopoietic stem and progenitor cells positively linked to glioblastoma progression. *Nat. Commun.* **12**, 1–16 (2021).
44. Pal, B. et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J.* **40**, e107333 (2021).
45. Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
46. Cheng, S. et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **184**, 792–809.e23 (2021).
47. Dundr, P. et al. Primary mucinous ovarian tumors vs. ovarian metastases from gastrointestinal tract, pancreas and biliary tree: a review of current problematics. *Diagn. Pathol.* **16**, 1–17 (2021).

48. Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. Hematopoiesis: A Human Perspective. *Cell Stem Cell* **10**, 120–136 (2012).
49. Liu, X. Y., Wu, J. & Zhou, Z. H. Exploratory undersampling for class-imbalance learning. *IEEE Trans. Syst. Man. Cybern. B. Cybern.* **39**, 539–550 (2009).
50. van Dijk, D. et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716–729.e27 (2018).
51. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R. N.* **2**, 18–22 (2002).
52. Ruedin, D. agrmt: Calculate Concentration and Dispersion in Ordered Rating Scales. R package version 1.42.8. https://CRAN.R-project.org/package=agrmt (2021).
53. Trang, N. V. et al. Determination of cut-off cycle threshold values in routine RT-PCR assays to assist differential diagnosis of norovirus in children hospitalized for acute gastroenteritis. *Epidemiol. Infect.* **143**, 3292–3299 (2015).
54. Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
55. Gamer, M., Lemon, J. & Singh, I. F. P. irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1. https://CRAN.R-project.org/package=irr (2019).
56. Quinn, T. peakRAM: Monitor the Total and Peak RAM Used by an Expression or Function. R package version 1.0.3. http://github.com/tpq/peakRAM (2022).
57. Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
58. Tickle, T. I., Georgescu, C., Brown, M. & Haas, B. inferCNV of the Trinity CTAT Project. https://github.com/broadinstitute/inferCNV (2019).
59. Dong, R. et al. Single-Cell Characterization of Malignant Phenotypes and Developmental Trajectories of Adrenal Neuroblastoma. *Cancer Cell* **38**, 716–733.e6 (2020).
60. Leader, A. M. et al. Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. *Cancer Cell* **39**, 1594–1609.e12 (2021).
61. Bi, K. et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell* **39**, 649–661.e5 (2021).
62. Couturier, C. P. et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat. Commun. 2020 111* **11**, 1–19 (2020).
63. Young, M. D. et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594–599 (2018).
64. Peng, J. et al. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res* **29**, 725–738 (2019). *2019 299*.
65. Nofech-Mozes, I., Soave, D., Awadalla, P. & Abelson, S. Data and Codes for Pan-cancer classification of single cells in the tumour microenvironment. https://doi.org/10.5281/zenodo.7419236 (2022).
66. Nofech-Mozes, I., Soave, D., Awadalla, P. & Abelson, S. abelson-lab/scATOMIC: scATOMIC v1.1.0. https://doi.org/10.5281/zenodo.7689011 (2023).

Acknowledgements

This work was supported by a grant from The Banting Research Foundation (Discovery Award #2021-1418) and Investigator Awards received from the Ontario Institute for Cancer Research with funds from the province of Ontario to S.A. and P.A. I.N.M. obtained funds from the Ontario Graduate Scholarship Program – University of Toronto. We thank Salman Basrai for his assistance checking the validity and accuracy of the scATOMIC code, as well as on his comments and suggestions concerning the package documentation.

Author contributions

All authors discussed the results, wrote, and commented on the paper. I.N.M. developed the concept, performed data analysis, developed the tool, and wrote the code. D.S. contributed to statistical analysis and model development. S.A. and P.A. conceived the idea, contributed to data analysis, led, and supervised all aspects of the study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-37353-8.

Correspondence and requests for materials should be addressed to Philip Awadalla or Sagi Abelson.

Peer review information *Nature Communications* thanks Maria Tsagiopoulou and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

nutrients

MDPI

Systematic Review

The Effect of Diet and Exercise Interventions on Body Composition in Liver Cirrhosis: A Systematic Review

Heidi E. Johnston [1,2,*], Tahnie G. Takefala [1], Jaimon T. Kelly [3,4], Shelley E. Keating [5], Jeff S. Coombes [5], Graeme A. Macdonald [2,6], Ingrid J. Hickman [1,2] and Hannah L. Mayr [1,2,7,8]

1 Department of Nutrition and Dietetics, Princess Alexandra Hospital, Woolloongabba, QLD 4102, Australia
2 Faculty of Medicine, The University of Queensland, Brisbane, QLD 4072, Australia
3 Centre for Online Health, Faculty of Medicine, The University of Queensland, Brisbane, QLD 4072, Australia
4 Centre for Health Services Research, Faculty of Medicine, The University of Queensland, Brisbane, QLD 4072, Australia
5 School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, QLD 4072, Australia
6 Department of Gastroenterology and Hepatology, Princess Alexandra Hospital, Woolloongabba, QLD 4102, Australia
7 Centre for Functioning and Health Research, Metro South Health, Brisbane, QLD 4102, Australia
8 Bond University Nutrition and Dietetics Research Group, Faculty of Health Sciences and Medicine, Bond University, Gold Coast, QLD 4226, Australia
* Correspondence: heidi.johnston@health.qld.gov.au; Tel.: +61-7-3176-7938

Abstract: Alterations in body composition, in particular sarcopenia and sarcopenic obesity, are complications of liver cirrhosis associated with adverse outcomes. This systematic review aimed to evaluate the effect of diet and/or exercise interventions on body composition (muscle or fat) in adults with cirrhosis. Five databases were searched from inception to November 2021. Controlled trials of diet and/or exercise reporting at least one body composition measure were included. Single-arm interventions were included if guideline-recommended measures were used (computed tomography/magnetic resonance imaging, dual-energy X-ray absorptiometry, bioelectrical impedance analysis, or ultrasound). A total of 22 controlled trials and 5 single-arm interventions were included. Study quality varied (moderate to high risk of bias), mainly due to lack of blinding. Generally, sample sizes were small (n = 6–120). Only one study targeted weight loss in an overweight population. When guideline-recommended measures of body composition were used, the largest improvements occurred with combined diet and exercise interventions. These mostly employed high protein diets with aerobic and or resistance exercises for at least 8 weeks. Benefits were also observed with supplementary branched-chain amino acids. While body composition in cirrhosis may improve with diet and exercise prescription, suitably powered RCTs of combined interventions, targeting overweight/obese populations, and using guideline-recommended body composition measures are needed to clarify if sarcopenia/sarcopenic obesity is modifiable in patients with cirrhosis.

Keywords: liver cirrhosis; sarcopenia; sarcopenic obesity; nutrition; exercise; body composition

Citation: Johnston, H.E.; Takefala, T.G.; Kelly, J.T.; Keating, S.E.; Coombes, J.S.; Macdonald, G.A.; Hickman, I.J.; Mayr, H.L. The Effect of Diet and Exercise Interventions on Body Composition in Liver Cirrhosis: A Systematic Review. *Nutrients* **2022**, *14*, 3365. https://doi.org/10.3390/nu14163365

Academic Editors: Aldo J. Montano-Loza and Maryam Ebadi

Received: 21 July 2022
Accepted: 12 August 2022
Published: 17 August 2022

1. Introduction

Advanced liver disease is a complex major health problem, impacting more than 1.5 billion individuals worldwide [1]. Cirrhosis is the end stage of chronic liver disease and is characterised by severe hepatic fibrosis with potential impacts on hepatic function. Once patients develop cirrhosis, they are at risk of dying from decompensated liver disease or hepatocellular carcinoma (HCC) [2]. Liver transplantation offers the opportunity to cure both. During the progression to cirrhosis, many aspects of health deteriorate, increasing the risk of malnutrition and loss of muscle mass [3,4], which in turn are associated with adverse outcomes for patients with cirrhosis and those awaiting transplant [5,6].

There are two key issues relating to body composition for people with liver cirrhosis. Firstly, sarcopenia is a condition characterised by a significant depletion of skeletal muscle in combination with low muscle strength and/or physical performance [7]. Sarcopenia is often interrelated with malnutrition [8]. In general, sarcopenia in liver disease literature refers to reduced muscle mass alone, which has a prevalence in cirrhosis between 40–70% [9]. Sarcopenia is associated with increased mortality in patients with cirrhosis, and in those who receive a liver transplant [10]. The second issue is an elevated body mass index in people with cirrhosis. Comorbid sarcopenia with obesity, where low muscle mass may be masked due to excess adiposity, increases the risk of hepatic decompensation and death in patients with cirrhosis [11,12]. Additionally, surgical risk is increased for obese liver transplant recipients [13,14]. The proportion of patients being referred for liver transplant with comorbid obesity is increasing [15]. Interventions to reduce adiposity may ameliorate the severity of their underlying liver disease, but also needs to be considered to improve transplant outcomes. The challenge in achieving weight loss in this patient group is to preserve or increase muscle mass whilst losing fat mass.

The first challenge in addressing low muscle mass and/or high adiposity in patients with cirrhosis is accurately assessing body composition, which can be complicated by fluid retention with ascites and oedema. Triceps skinfold thickness (TSF) and mid-arm muscle circumference (MAMC) appear less affected by fluid overload than other anthropometric measures in this population [16]. While there is evidence that these measures have good intra- and inter-rater reliability for the diagnosis of malnutrition [17], there remain concerns about their reproducibility [18] and their reliability in identifying subtle changes [7]. Recent guidelines have recommended several reference methods for the assessment of body composition in patients with cirrhosis, specifically computerised tomography (CT) and magnetic resonance imaging (MRI) techniques [16,19]. The use of dual-energy Xray absorptiometry (DXA), bioelectrical impedance analysis (BIA), and ultrasound are also supported when CT/MRI are unavailable. These may be more readily available in clinical settings, although the reliability of BIA and some DXA measures may be adversely impacted by fluid retention in decompensated cirrhosis [20,21].

It is well known that both diet and exercise have positive effects on health outcomes across multiple chronic health conditions. Low muscle mass and high adiposity are attractive therapeutic targets in advanced liver disease because they may be modifiable through diet and/or exercise interventions. Exercise training is known to reduce the progression, or reverse muscle wasting [22] and has been shown to improve physical function and frailty in cirrhosis [23]. According to current guidelines [16,19], a high protein, high energy diet has been recommended for people with cirrhosis, due to catabolic effects of cirrhosis that can lead to protein degradation and therefore muscle loss. Minimising fasting times, and the inclusion of a late evening carbohydrate rich snack to prevent overnight catabolism have also been recommended [24]. It is still unclear how to accurately estimate energy needs for individuals with cirrhosis who are obese. There have also been several studies exploring the effect of Branched Chain Amino Acids (BCAAs) in this population; however, there has been heterogeneity in BCAA dosage type [25]. While advice about combined diet and exercise in cirrhosis is beginning to appear in more detail in practice guidelines [26], there remains a gap in current knowledge relating to improving body composition in cirrhosis, especially in obese persons. There is currently no comprehensive synthesis of evidence to guide interventions to slow progression or potentially reverse muscle wasting or reduce adiposity for patients with cirrhosis.

Therefore, we aimed to systematically evaluate the evidence on the effect of diet and/or exercise interventions on body composition in adults with cirrhosis, with a particular interest in the impact of these interventions on patients with obesity and liver cirrhosis to determine whether muscle mass can be preserved concurrently with fat loss.

2. Materials and Methods

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement [27] (see Supplementary Materials Supplementary File S1), and the protocol was registered with the international Prospective Register of Systematic Reviews (PROSPERO ID: CRD42020176547).

2.1. Eligibility Criteria

Table 1 summarises the population, intervention, control, outcomes, and study design (PICOS) for the study selection.

Table 1. PICOS for study selection and eligibility criteria.

Criteria	Inclusion and Exclusion Details
Population	- Liver cirrhosis, including potential transplant candidates.
Intervention	- Diet or exercise intervention (alone or combination), of at least four weeks duration. - Studies excluded if the intervention was a single nutrient (e.g., vitamin D, omega-3 fatty acid), or nutrition was exclusively administered intravenously without oral nutrition support.
Control	- No specified control. - Studies without a control group were included if they reported specific body composition measures (see below).
Outcomes	- At least one body composition measure, via imaging (CT, MRI, or DXA), BIA, ultrasound, or anthropometry (TSF, MAMC, MAC, thigh, or calf circumference). - Single-arm interventions were included if they had one of the guideline-recommended measures (CT, MRI, DXA [19]; or BIA if in compensated cirrhosis). - Waist circumference was not included due to the confounding effect of any ascites.
Study Design	- RCTs, non-randomised controlled trials and single-arm interventions were eligible. - Articles excluded: case report, letter to the editor, abstract only, or non-English.

CT: computerised tomography, MRI: magnetic resonance imaging, DXA: dual-energy Xray absorptiometry, BIA: bioelectrical impedance analysis, TSF: triceps skinfold thickness, MAMC: mid-arm muscle circumference, MAC: mid-arm circumference, RCT: randomised controlled trials.

2.2. Search Strategy

Databases were searched from inception to 15 November 2021 (PubMed, Embase, Web of Science, CINAHL, and CENTRAL). Reference lists of relevant review articles were hand-searched to identify further articles. The strategy utilised a combination of key words and controlled vocabulary combining terms related to liver cirrhosis AND diet/exercise AND intervention/trial (see Supplementary Materials Supplementary File S2 for full search strategy). The final search was de-duplicated using reference management software, Endnote [28]. References were screened in Rayyan [29]. Two reviewers (H.J. and T.T.) independently screened approximately half of the title and abstracts using a screening tool. Twenty studies were piloted with the tool to determine agreement before completing the screening. For potentially eligible articles, full texts were retrieved and independently screened by two of three reviewers (H.J., T.T., or H.M.). Disagreements were resolved by consensus or referral to the third reviewer.

2.3. Data Extraction

Extracted data included study authors, publication year, country, population, setting, intervention, control, and body composition outcomes. If data were not available, an attempt to contact authors was made to retrieve information. Data were initially extracted by either of two reviewers (H.J. or T.T.) in a standardised extraction table. Extraction was piloted across three different study designs (RCT, non-randomised controlled trial, and single-arm intervention studies) to ensure consistency. Where present, we extracted body composition change data between study treatment groups. If unavailable, we recorded the within-group change. All extraction was cross-checked by a second reviewer with disagreements discussed to reach consensus.

2.4. Quality Assessment

For each included study, risk of bias was assessed independently by two of three reviewers (H.J., T.T., or H.M.) using the Cochrane risk of bias tool (Rob2) [30] for RCTs, and the ROBINS-I tool [31] for non-randomised controlled and single-arm studies. Rob2 evaluates five domains including risk of bias from: randomisation process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. The ROBINS-I tool evaluates seven domains including risk of bias due to: confounding, participant selection, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of the reported result. For the domains considering deviations from intended interventions, where intervention blinding is considered, we allocated 'some concerns' rather than 'serious' if participants were not blinded. This is due to the nature of diet and/or exercise interventions, where it is often not feasible for intervention allocation blinding. Conflicts were resolved by consensus or a third reviewer. The certainty of the body of evidence based on outcomes using the Grading of Recommendations Assessment Development and Evaluation was not possible due to significant variability in study design, interventions used, outcome measures employed, and statistical methodologies used across the studies.

2.5. Data Synthesis

A meta-analysis was unable to be performed due to variability in study interventions, control groups, tools to assess body composition, and reporting of means and medians across studies. Narrative synthesis was conducted based on type of intervention and body composition measures. Where a study reported on multiple body composition measures the guideline-recommended measures were prioritised in the text results (CT or MRI, followed by DXA, BIA, or ultrasound, then anthropometry).

3. Results

3.1. Characteristics of Studies

The final search contained 10,099 articles, including three articles from hand searches (Figure 1). A total of 152 full text articles were retrieved and 27 studies included in this review. Thirty-two studies were excluded for not reporting on body composition measures. The characteristics and outcomes of the included studies are summarised in Table 2. Of the 27 studies, 19 were RCTs, 3 were non-randomised controlled trials, and 5 were single-arm intervention studies. Most studies were relatively small, with participant numbers ranging from 6 to 120, totalling 1263 participants. Intervention duration ranged between 4 and 56 weeks and populations included patients with both compensated and decompensated cirrhosis. Only one study specifically targeted an overweight population [32], however the primary outcomes of interest were weight loss and portal hypertension changes. Thirteen studies in total reported populations with a mean BMI either overweight [33–40] or obese [32,41–44]. Others did not report on BMI [45–51]. None of the studies specifically targeted sarcopenic obesity in cirrhosis.

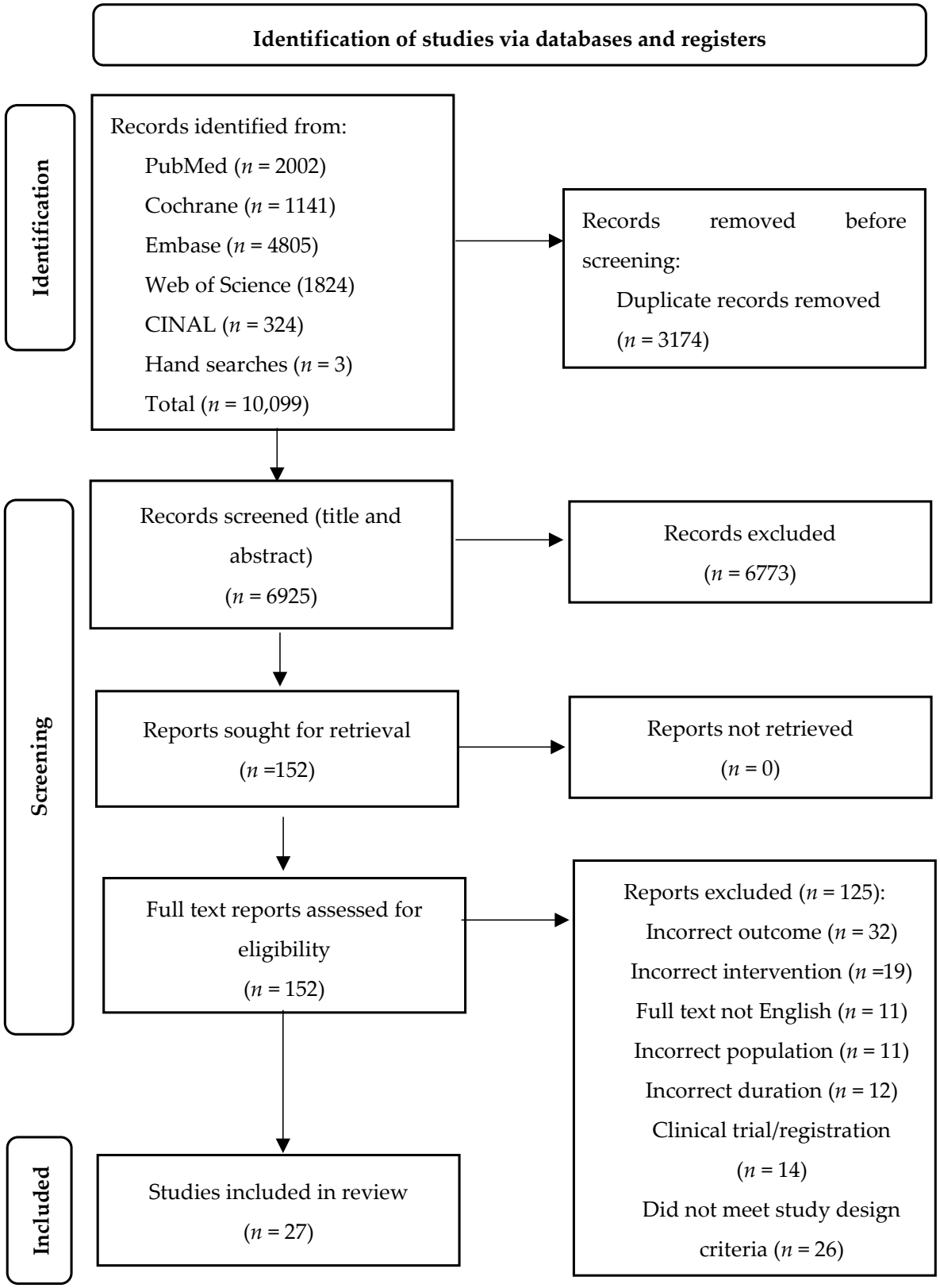


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the study selection process.

Table 2. Study characteristics and outcomes for diet and/or exercise interventions in cirrhosis.

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Combined intervention studies (n = 9 RCTs, n = 2 non-randomised studies, n = 3 single arm intervention trials)						
Aaman et al. [40] 2019 Denmark	RCT	Intervention *n* = 20 Age 61.7 ± 7.8 years 80% male BMI 26 ± 3.0 kg/m^2 Child Pugh Class: A 50%, B 50% MELD 10.8 ± 2.7 Control *n* = 19 Age 63 ± 7 years 74% male BMI 25 ± 4.2 kg/m^2 Child Pugh Class: A 53%, B 47% MELD 10.7 ± 2.8 Outpatients	Supervised resistance training 3 days/week for 60 min at a moderate level. 5 min warm up, then 7 whole body exercises, (3 sets for legs, 2 sets for arms/chest, 1 set lower back, 1 for abdominals), starting at 15–12 repetitions at the start down to 8 by week 12 Duration:12 weeks	Oral nutrition supplements (125 mL, 14.4 g protein and 2.9 g BCAA/100 g) provided if protein intake < 1.2 g/kg/day at baseline	No change to current exercise or diet	*Intervention versus control:* ↑ Cross sectional area of quadriceps via MRI ↑ Body cell mass via BIA ↔ Dry lean mass via BIA ↔ Lean mass via BIA ↔ Calf circumference ↔ MAC ↔ Thigh circumference ↔ Mid arm muscle area ↔ TSF
Chen et al. [44] 2020 USA	Pilot RCT	Intervention *n* = 9 Age 55 ± 7 years 56% male BMI 30 ± 6 kg/m^2 Child Pugh Class: B 78%, C 22% MELD-Na 16 ± 4 Control *n* = 8 Age 54 ± 11 years 75% male BMI 31 ± 8 kg/m^2 Child Pugh Class: B 50%, C 50% MELD-Na 19 ±3 Portal hypertension and MELD ≥ 10 Outpatients	Education on exercise, and behavioural counselling bi-weekly for first 8 weeks. Self-directed exercise increasing 500 steps/day weekly to biweekly. Daily to weekly motivational phone calls. Duration: 12 weeks	Standardised diet provided 1.2–1.5 g/kg/day of protein + late evening snack + oral nutrition supplement (6 g essential amino acids) twice a day	Standardised diet (same as intervention group) only	*Intervention versus control:* ↑ Psoas muscle index via CT ↔ Total skeletal muscle index via CT ↔ Intramuscular adipose tissue via CT ↔ Total abdominal adipose tissue via CT ↔ Total thigh muscle volume via CT ↔ Thigh muscle index via CT ↔ Cross sectional area, 50% of femur length via CT ↔ Thigh adipose tissue volume via CT ↔ Fat free mass via DXA ↔ Fat mass via DXA ↔ Lean muscle index via DXA ↔ Lower extremities lean muscle index via DXA ↔ Fat free mass via BIA ↔ Fat mass via BIA ↔ Skeletal muscle mass via BIA ↔ Skeletal muscle index via BIA ↔ Phase angle via BIA

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or vs. Control)*
Hernandez-Conde et al. [39] 2021	Pilot, double-blind RCT	Intervention n = 15 Age 69 ± 9.7 years 86.7% male BMI 29 ± 4.6 kg/m^2 MELD 10.7 ± 4.4 Child Pugh Class: A 78.6%, B 21.4% Control n = 17 Age 61 ± 9.4 years 88.2% male BMI 26 ± 4.7 kg/m^2 MELD 11 ± 3.4 Child Pugh Class: A 59%, B 29%, C 12% Compensated outpatients	Personalised exercise instructions with use of accelerometers in wristbands or smartphones to include 5000–10,000 steps/day with gradual increments of 2000–2500 steps/day + moderate intensity exercise in 30-min sessions (goal at least 150 min/week) + verbal reinforcement at reviews. Duration: 12 weeks	Personalised diet recommendations + instructed to eat 7 meals/day including late evening snack plus BCAA supplement 100 g dissolved in 500 mL water throughout the day (15 g protein, 8.5 g fat, 68 g of carbohydrates, 2.61 g of leucine, 1.01 g of isoleucine, and 1.62 g of valine) + verbal reinforcement at reviews	Same exercise and diet recommendations as intervention group except took placebo supplement 100 g dissolved in 500 mL water throughout day (maltodextrin 99.63%) instead of BCAA	*Intervention versus control:* ↑ Skeletal muscle index via CT ↓ % total body fat via BIA ↔ Phase angle via BIA
Kruger et al. [47] 2018 Canada	RCT	Intervention n = 20 Age 53 ± 8 years 50% male MELD 9.05 Child Pugh Class: A 70%, B 30% Control n = 18 Age 56.4 ± 8.5 years 65% male MELD 9.7 Child Pugh Class: A 70%, B 30% BMI not reported Outpatients	Supervised at home, moderate to high intensity aerobic exercise (60–80% of heart rate reserve) on cycle ergometer 3 days/week (30 min sessions gradually increased to 60 min). Visited bi-weekly for session observation. Duration: 8 weeks	Dietary counselling on optimal protein (1.2–1.5 g/kg/day, ideal body weight for BMI > 30) and energy intake (35–40 kcal/kg for BMI 20–30, 25–35 kcal/kg for BMI 30–40, and 20–25 kcal/kg for BMI > 40. Advised on exercise days to consume an extra 250–300 kcal.	Usual care	*Intervention versus control:* ↔ Thigh muscle mass via ultrasound ↔ Thigh circumference

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Lattanzi et al. [38] 2021	Pilot single blind RCT	Intervention n = 14 Age: 59.2 ± 8.4 years 64% male BMI 29.8 ± 4.3 kg/m^2 Child Pugh Class: A 86%, B 14% MELD 9 ± 2.7 Control n = 10 Age: 56 ± 4.6 years 60% male BMI 29.6 ± 6.8 kg/m^2 Child Pugh Class: A 90%, B 10% MELD 9.8 ± 3.2 Outpatients with portal hypertension	Motivational interviewing with information on physical activity at baseline	Motivational interview at baseline with information and counselling on diet in line with EASL clinical guidelines (2019) + HMB supplement (3 g/day)	Same exercise and diet as intervention group + placebo supplement (Sorbitol 3 g/day)	*Within group changes:* ↑ Thigh muscle thickness via ultrasound ↔Fat free mass via BIA ↔Phase Angle via BIA
Macias-Rodriguez et al. [37] 2020	RCT	Intervention n = 22 Age 53.5 ± 7.6 years 47% male BMI 29.8 ± 4.8 kg/m^2 Child Pugh Class: A 82%, B 18% MELD 8.5 (7–10) Control n = 21 Age 53.7 ± 8.2 years 43% male BMI 29.2 ± 3.7 kg/m^2 Child Pugh Class: A 95%, B 5% MELD 8 (7.5–9.5) Compensated cirrhosis, outpatients	Given wrist-worn accelerometer as activity tracker. Aim to gradually increase physical activity to reach >2500 steps/day above baseline. Total 5000 steps/day. Light to moderate intensity. Duration: 10 weeks	Harris–Benedict equation was utilised to calculate energy requirements + 10% extra for thermic effect of food and 20% extra for exercise. Diet 60% carbohydrates, 1.3–1.5 g protein/kg/day + remainder from fats + 1.5–2 g sodium restriction/day restriction + non-alcoholic beer at lunch (330 mL/day)	The same diet and exercise prescribed as intervention group without non-alcoholic beer (given a 330 mL bottle of water instead)	*Within group changes:* ↔ Phase Angle via BIA ↑ Thigh circumference ↔MAMC ↔TSF

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or vs. Control)*
Macias-Rodriguez et al. [36] 2016 Mexico	Pilot open RCT	Intervention *n* = 13 Age 53 (48–55) years 69% male BMI 27.5 (22.4–28.9) kg/m^2 Child Pugh score 6 (5–7) MELD 9 (8–12) Control *n* = 12 Age 51 (38–57) years 83% male BMI 27.4 (25–30) kg/m^2 Child Pugh score 6 (5–7) MELD: 12 (7–14) Compensated outpatients	Supervised exercise 3 days/week of 60–70% max heart rate, for 40 min of aerobic training using cycle ergometer + kinesiotherapy/rhythmic activities) Duration: 14 weeks	Instructed to consume 30% extra calories (65% carbohydrates, 1.2 g/kg/day protein) + no added salt diet of 1.5–2 g/day	Same recommendations as intervention; consume 10% extra calories (65% carbohydrates, 1.2 g/kg/day protein) + no added salt diet of 1.5–2 g/day. Continue regular activities, no new exercise	*Intervention versus control:* ↑ Phase angle via BIA
Roman et al. [33] 2014 Spain	Pilot RCT	Intervention *n* = 8 Age 65.5 (46–72) years 62% male BMI 26.7 (18.3–34.7) kg/m^2 Child Pugh Class: A 87%, B 13% MELD 9.5 (7–12) Control *n* = 9 Age 61 (43–75) years 78% male BMI 27.6 (19.5–35.3) kg/m^2 Child Pugh Class: A 78%, B 22%, MELD 9 (7–13) Outpatients with a previous episode of decompensation	Supervised exercise 3 days/week, moderate intensity (60–70% max heart rate) for 60 min. Cycle ergometry and treadmill walking Duration: 12 weeks	10 g oral leucine supplementation daily	10 g oral leucine supplementation daily, no exercise recommendations	*Within group changes:* ↑ Lower thigh circumference (intervention compared to baseline, ↔ control) ↔ Mid or upper thigh circumference (intervention or control) ↔ MAMC (intervention or control) ↔ Mid-arm circumference (intervention or control) ↔ TSF (intervention or control)

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or vs. Control)*
Zenith et al. [35] 2014 Canada	RCT	Intervention n = 9 Age 56 ± 8 years 78% male BMI 27.7 ± 3.8 kg/m^2 Child Pugh score: 6.2 ± 1.4 MELD 9.7 ± 2.4 Control n = 10 Age 59 ± 6 years 80% male BMI 28.9 ± 4.1 kg/m^2 Child Pugh score: 6.3 ± 1.4 MELD 10.2 ± 1.9 Outpatients, Child Pugh A or B	Supervised exercise 3 days/week, 60–80% of peak VO_2, 30 min session, increased by 2.5 min per session each week, 5 min warm up and cool down using cycle ergometer Duration: 8 weeks	Baseline dietetic counselling to reach 1.2–1.5 g/kg of protein (for BMI > 30 adjustments made based on ideal body weight), calories BMI specific (between 14 up to 30 kcal/kg) and instructed to consume an extra 250–300 calories on exercise days	Baseline counselling by dietitian (same as intervention) but no formal exercise regimen	*Intervention versus control:* ↑ Quadricep muscle thickness via ultrasound ↑ Thigh circumference
Morkane et al. [43] 2020 United Kingdom	Non-randomised controlled trial	Intervention n = 16 Age 55.6 ± 7.8 years 87.5% male MELD 13.7 ± 4.6 BMI 30.9 ± 5.6 kg/m^2 Control n = 17 Age 55.6 ± 7.8 years 82.7% Male MELD 13.2 ± 3.7 BMI: 27 ± 4.6 kg/m^2 Outpatients, transplant candidates	Supervised 40 min interval training on cycle ergometer (4–6 × 3 min intervals at 80% of AT (moderate intensity) and 4–6 × 2 min intervals at 50% of difference between VO_2 at peak and VO_2 at AT ('severe' intensity) with 5 min warm up and cool down) Duration: 6 weeks	Standardised nutrition assessment and advice by transplant dietitian at baseline and 6 weeks	Standard care, no initiation of exercise. Standardised nutrition assessment and advice by transplant dietitian at baseline and 6 weeks	*Within group changes:* ↔ Mid-arm circumference (intervention or control) ↔ MAMC (intervention or control)

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or vs. Control)*
Schmidt et al. [42] 2021	Non-randomised controlled trial	Intervention n = 11 Age 56.6 ± 9.9 years 63.6% male BMI 30.3 ± 5.4 kg/m^2 Child Pugh Class: A 91%, B 9% Control n = 22 Age 58.7 ± 12.9 years 59.1% male BMI 32.4 ± 5.1 kg/m^2 Child Pugh Class: A 86%, B 14% MELD—not reported Compensated outpatients	Supervised exercise 3 days/week, aerobic, moderate intensity (5 min warm up, 30 min walking/running 60–70% VO_2 max). Increasing session by 2 min until reaching 50 mins by week 8. Duration: 12 weeks	Diet advice to aim for 25–30 kcal/day and 1.2–1.5 g of protein/kg/day—using estimated dry body weight.	The same diet advice without any exercise intervention	*Intervention versus control:* ↔ Phase Angle via BIA ↔ Lean mass via BIA ↔ Fat mass via BIA ↔ MAMC ↓ MAC
Berzigotti et al. [32] 2017 Spain	Multi-centre single arm intervention pilot study	Total n = 50 Age 56 ± 8 years 62% male BMI 33.3 ± 3.2 kg/m^2 MELD 9 ± 3 Child Pugh Class: A 92%, B 8% Compensated outpatients with BMI ≥ 26 kg/m^2	Supervised exercise 1 day/week for 60 min moderate intensity (10–12 Borg Scale of Perceived Effort) in groups of 1–5 + increase daily step activity Duration: 16 weeks	Reduction of 500–1000 kcal/day. Protein intake maintained at 20–50% of total kcal and within 0.8 g/kg ideal bodyweight/day. Carbohydrates 45–50% and fat <35% of total kcal. 20 g/day alimentary fibre recommended.	No control	↓ Fat mass via BIA ↔ Lean mass via BIA
Hiraoka et al. [52] 2017 Japan	Single arm inter-vention study	Total n = 33 Age 67 (63–71) years 39% men BMI 23.2 (20.8–25.1) kg/m^2 Child Pugh Class: A 90%, B 10% Compensated outpatients	Walking (an additional 2000 steps/day on top of usual average steps) Duration: 12 weeks	Late evening BCAA supplement provided once daily (13.5 g protein, 210 kcal/day)	No control	↑ Muscle volume via BIA (reported as change ratio)

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Nishida et al. [53] 2017 Japan	Single arm intervention study	Total *n* = 6 Age from 51–79 years 100% female BMI 24.3 (19.6–26.1) kg/m^2 Child Pugh Class: A 100% Compensated outpatients	Instructed to undertake bench step activity at anaerobic threshold level at home. Aim 140 min/week. Duration: 12 months	BCAA supplement (3 sachets/day = 12.45 g of BCAA), no specific nutrition advice except to maintain usual dietary intake	No control	↔ % fat via BIA ↔ Visceral fat area via CT ↔ Intramuscular adipose tissue content via CT
Diet-only intervention studies (n = 9 RCTs, n = 1 non-randomised study, n = 2 single arm interventions)						
Dupont et al. [45] 2012 France	Multi-centre RCT	Intervention *n* = 44 Age 56.1 ± 9.6 years 68% male Child Pugh score: 11.2 ± 1.3 Control *n* = 55 Age 54.6 ± 9.6 years 64% male Child Pugh score: 10.5 ± 1.5 BMI or MELD—not reported Inpatients with ARLD and jaundice (without alcoholic hepatitis)	NA	Enteral nutrition 3–4 weeks (30–55 kcal/kg/day through nasogastric tube). Subsequent 3 oral nutrition supplements/day for 2 months Duration: 12 weeks with outcomes reported at 12 months	Standard hospital oral diet	*Intervention versus control:* ↔ MAMC ↔ TSF
Hirsh et al. [54] 1983 Chile	RCT	Intervention *n* = 26 Age 49.9 ± 8.6 years 81% male Control *n* = 25 Age 46.1 ± 8.0 years 84% male BMI, Child Pugh, or MELD—not reported Decompensated outpatients	NA	1 L oral nutrition supplement /day (1000 kcal, 34 g protein) + usual diet Duration: 12 months	Placebo tablet daily	*Intervention versus control:* ↔ TSF ↔ Mid-arm circumference

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Le Cornu et al. [55] 2000 England	RCT	Intervention n = 42 Age 52 (27–67) years 69% male Child Pugh Class: A 7%, B 48%, C 45% Control n = 40 Age 50 (24–68) years 78% male Child Pugh Class: A 10%, B 28%, C 62% BMI or MELD not reported Outpatient transplant candidates with MAMC < 25% percentile	NA	Oral nutrition supplement of 500 mL/day (750 kcal, 20 g protein) was given + dietary counselling to adapt/increase their calories and protein based on their medical condition until transplantation Duration: until transplantation. Median wait 77 (1–395) days intervention and 45 (1–424) control	Standard dietary advice to adapt/increase their calories and protein based on their medical condition until transplantation	*Intervention versus control:* ↔ MAMC ↔ Mid-arm circumference ↔ TSF
Les et al. [48] 2011 Spain	Multi-centre RCT	Intervention n = 58 Age 64.1 ± 10.4 years 78% male Child Pugh 8.3 ± 2.0 MELD 16.1 ± 4.5 Control n = 58 Age 62.5 ± 10.4 years 74% male Child Pugh 8.1 ± 1.7 MELD 16.2 ± 3.9 BMI—not reported Outpatients with previous episode of hepatic encephalopathy	NA	Diet of 35 kcal/kg + 0.7 g/kg of protein/day adjusted to ideal weight + late evening BCAA supplement 2/day (120 kcal). Enteral nutrition if admitted for episode of hepatic encephalopathy and oral intake in hospital was poor. Duration: mean 32 ± 22 weeks intervention and 36 ± 2 weeks control	Same diet but with maltodextrin supplement 2/day instead of BCAA. Enteral nutrition provided if episode of hepatic encephalopathy and oral intake was poor	*Within group changes:* ↑ MAMC (intervention compared to baseline) ↔ MAMC (control compared to baseline)

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Manguso et al. [34] 2005 Italy	Random-ised, double period cross-over trial	Group 1: n = 45 Age 60 ± 9 years 67% male BMI 28.5 ± 3.2 kg/m^2 Child Pugh Class: A 33%, B 77% Group 2: n = 45 Age 60 ± 7 years 49% male BMI 27.8 ± 2.1 kg/m^2 Child Pugh Class: A 33%, B 77% Outpatients with HCV cirrhosis	NA	Group 1: Prescribed diet of 30–40 kcal/kg/day based on calculated desirable weight (total calories split into 16% protein, 55% carbohydrates, 28–30% fat) + low sodium 1000 mg/day Followed by usual diet after. Group 2: Usual diet first. Followed by prescribed diet second. Duration: 3 months per diet (6 months total)		*Within group changes:* ↑ MAMC (Group 1 at 3 months post prescribed diet vs baseline) ↑ MAMC (Group 2 at 6/12, post prescribed diet vs baseline and vs 3/12) ↓ MAMC (Group 1 at 6 months post usual diet vs 3 months post prescribed diet) ↔ MAMC (Group 2 at 3 months post usual diet vs baseline) ↔ TSF (Group 1 or Group 2 after both diet interventions at 3 and 6 months)
Okabayashi et al. [56] 2011 Japan	RCT	Intervention n = 40 Age 68 ± 7.6 years 28% male BMI 23.6 ± 3.2 kg/m^2 Child Pugh Class: A 70%, B 30% Control n = 36 Age 65.1 ± 11.3 years 31% male BMI 22.7 ± 3.2 kg/m^2 Child Pugh Class: A 71%, B 29% Outpatients with scheduled HCC surgery	NA	Carbohydrate and BCAA enriched supplement morning and night. (420 kcal, 13 g free amino acids, 13 g of gelatine hydrolysate, 62 g carbohydrates, 7 g lipids) + dietitian education to modify intake to reduce 420 kcal/day to account for the supplement and match caloric intake to controls Duration: supplements for at least 6 months, with a follow up at 12 months	Usual diet. No supplements	*Intervention versus control:* ↑ MAMC (at 6, 8, 10, 12 months) ↔ TSF no change post-operatively in both groups (data not reported)

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Poon et al. [50] 2004 China	RCT	Intervention n = 41 Age 59 (24–84) years 95% male Control n = 43 Age 59 (27–80) years 90% male. No BMI, Child Pugh or MELD reported. Outpatients with unresectable HCC	NA	BCAA supplement morning and night (420 kcal, 13 g amino acids, 13 g peptides, 62 g carbohydrates, 7 g lipids) + unrestricted diet unless HE—protein was restricted Duration: 1 week prior to surgery, up to 12 months	Usual diet	*Intervention versus control:* ↔ Mid-arm circumference ↔ TSF
Sorrentino et al. [51] 2012 Italy	RCT	Group A: n = 40 Age 64 ± 6.3 years 65% male Child Pugh Class: B 28%, C 72% MELD 12.1 ± 0.7 Group B: n = 40 Age 66 ± 7.5 years 67% male Child Pugh Class: B 30%, C 70% MELD 11.7 ± 0.7 Group C: n = 40 Age: 65 ± 7.6 years 70% male Child Pugh Class: B 25%, C 75% MELD 12.4 ± 0.9 BMI not reported In/outpatients with refractory ascites	NA	Group A: Instructed to consume 1–1.3 g protein/kg/day, 30–35 kcal/kg/day + low sodium diet (80 mEq/day) + BCAA evening snack (210 kcal, 13.5 g protein, 3.5 g fat) + instructed to adjust energy intake to account for BCAA supplement + post LVP parenteral nutrition for 24 hrs post paracentesis during hospital admission + Dietitian advice monthly. Group B: same as group A without parenteral nutrition post paracentesis. Duration: 12 months, follow up at 3, 6, 12 months	Group C: Low sodium diet (80 mEq /day) + Dietitian advice monthly	*Between group changes:* ↓ TSF (Group C versus Group A at 3, 6, and 12 months and Group C versus Group B at 6 months only) ↓ MAC (Group C versus Group A and Group B at 6 and 12 months)

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or vs. Control)*
Tangkijvanich et al. [57] 2000 Thailand	RCT	Group 1: n = 14 Age: 53 ± 11 years 71% male BMI 23.7 ± 3.4 kg/m^2 Child Pugh score: 5–7: 64%, score 8–15: 36%. Group 2: n = 15 Age: 53 ± 13 years 80% male BMI: 25 ± 4.1 kg/m^2 Child Pugh score: 5–7: 60%, score 8–15: 40% Outpatients	NA	Group 1: received standard diet (40 g protein/day) + 150 g BCAA supplement/day = total of ~2000 kcal/day. Duration: 4 weeks	Group 2: standard diet (80 g protein/day = total of ~2000 kcal/day)	*Within group changes:* ↔ MAMC (Group 1 or Group 2)
Okabayashi et al. [49] 2008 Japan	Non-randomised study with historical control group	Intervention n = 13 Age 66.2 ± 9.1 years 54% male Child Pugh Class: A 77%, B 23% Control n = 28 Age 65.6 ± 8.2 yrs 75% male Child Pugh Class: A 82%, B 18% BMI not reported Outpatients for HCC surgery	NA	Carbohydrate and BCAA enriched supplement morning and night. (420 kcal, 13 g free amino acids, 13 g gelatin hydrolysate, 62 g carbohydrates, 7 g lipids) Duration: 2 weeks prior to surgery and at least 6 months post	Usual care—no supplementation	*Within group changes:* ↑ MAMC (baseline to 6 months for intervention, not reported for control)
Kitajima et al. [59] 2018 Japan	Single arm intervention study	Total n = 21 Age 71.3 ± 7.9 years 42% male BMI 23.9 ± 4.0 kg/m^2 Child Pugh Class: A 48%, B 52% MELD—not reported Outpatients with hypoalbuminaemia	NA	BCAA supplement 3/day after meals. Dietitian advised intakes of 25–35 kcal/kg/day and protein 1–1.4 g/kg/day. Adherence monitored monthly. Duration: 48 weeks	No control	↔ Skeletal muscle index via CT ↔ Intramuscular adipose tissue content via CT ↔ Subcutaneous fat area via CT ↔ Visceral fat area via CT

Table 2. *Cont.*

Study Citation, Country	Study Design	Population	Exercise Intervention	Dietary Intervention	Control Group	Body Composition Outcomes ↑ = *Significantly Increased or Higher* ↓ = *Significantly Decreased or Lower* ↔ = *No Significant Difference (Pre/Post or* vs. *Control)*
Putadechakum et al. [58] 2012 Thailand	Single arm intervention study	n = 22 Age 52.9 ± 12.8 years 55% male BMI 21.4 ± 0.6 kg/m^2 Child Pugh Class: A 63%, B 23%, C 14% Outpatients with ARLD	NA	20 g protein (soy based) oral nutrition supplement daily (420 kcal, 20 g protein, 65 g CHO, 10.6 g fat) + regular diet. Duration: 8 weeks	No control	↑ Lean mass via BIA ↔ Fat mass via BIA ↔ TSF
Exercise only intervention (n = 1 RCT)						
Roman et al. [41] 2016 Spain	RCT	Intervention n = 14 Age 62 ± 2.4 years 71% male BMI 31.5 ± 1.6 kg/m^2 Child Pugh score: 5.4 ± 0.2 MELD 8.2 ± 0.4 Control n = 9 Age 63.1 ± 2.3 years 85% male BMI 30.3 ± 1.4 kg/m^2 Child Pugh score: 5.4 ± 0.2 MELD 9.1 ± 0.4 Outpatients with a previous episode of decompensation	Supervised exercise 3 days/week, 60 min of cycle ergometry and treadmill walking + 5–10 min of upper body resistance exercise + 10–15 min balance, coordination, stretching and relaxation. Moderate intensity (60–70%) of max heart rate. Duration: 12 weeks	NA	Sham intervention 1 h 3 days/week of cephalocaudal muscle relaxation, and breathing, visualisation, and concentration exercises	*Within group changes:* ↑ Lean appendicular mass via DXA (intervention compared to baseline, ↔ control) ↑ Lean leg mass via DXA (intervention compared to baseline, ↔ control) ↑ Lean body mass via DXA (intervention compared to baseline, ↔ control) ↓ Fat body mass via DXA (intervention compared to baseline, ↔ control) ↑ Upper thigh circumference (intervention compared to baseline, ↔ control) ↔ Lower thigh circumference (intervention or control) ↓ Mid-arm circumference and mid-arm skinfold thickness (intervention compared to baseline, ↔ control) ↓ Mid-thigh skinfold thickness (intervention compared to baseline, ↔ control) ↔ MAMC (intervention or control)

Outcome data presented for controlled trials are the between group differences (where reported) and the within group differences if the significance of between group data were not reported. Data presented as mean SD or median (range/inter-quartile range). RCT: randomised controlled trial, AT: anaerobic threshold, MELD: model for end-stage liver disease, BMI: body mass index, ARLD: alcohol related liver disease, BCAA: branched-chain amino acid, CT: computed tomography, DXA: dual-energy X-ray absorptiometry, BIA: bio-electrical impedance analysis, MAMC: mid-arm muscle circumference, TSF: triceps skinfold thickness, MAC: mid arm circumference, HE: hepatic encephalopathy, LVP: large volume paracentesis, EASL; European Association of the Study of the Liver, NA: not applicable, VO_2 max: maximum amount of oxygen your body is able to use during exercise. Child Pugh score [60].

For most of the included studies, the change in muscle or fat mass was a secondary outcome, and factors such as muscle strength [29], aerobic/exercise capacity [41,46], survival [39,47], quality of life [48], portal hypertension [28], hepatic venous pressure [30], or liver function [49] were primary outcomes. Fourteen studies were combined diet and exercise interventions [32,33,35–40,42–44,47,52,53], all in outpatient settings. Their exercise components varied, with most delivering supervised sessions in a clinic setting [32,33,35,36,40,42,43], although one study was supervised by a clinician at the patient's home [47] and others were self-directed at home [37–39,44,52,53]. Most exercise was moderate to high intensity for 30–60 min sessions on 1 to 3 days a week and utilised either aerobic or resistance training, or a combination of these. Otherwise, some self-directed sessions focused on increasing step counts. The dietary component of five of these combined interventions used a high protein and energy diet [35,36,40,42,44]. Four of those studies provided the same dietary intervention to the control group, with the only difference between treatment groups being exercise in the intervention arms [35,36,42,44], while only one study provided "usual care" to the control participants [40]. Another combined study followed this style, however providing 'standard dietary advice' to both intervention and control arms, while the intervention arm also received supervised exercise training [43]. Three other combined studies delivered exercise and diet interventions to both groups, with the difference being a specific dietary product, either non-alcoholic beer [37], branched-chain amino acids (BCAAs) [33,39], or beta-hydroxy-beta-methylbutyrate (HMB, a metabolite from leucine) [38]. Of the three diet and exercise single-arm interventions, two provided BCAAs, with self-directed exercise [52,53] while the third study in overweight cirrhotic patients focused on a hypocaloric, moderate protein diet with supervised exercise [32].

Twelve diet-only studies [34,45,48–51,54–59] were included: ten in an outpatient setting [34,48–50,54–59], one in inpatients [45], and the twelfth commenced in inpatients with outpatient follow up [51]. Most (n = 9) interventions prescribed a high protein and energy diet plus oral nutritional supplements either with [48–50,56,57,59] or without BCAAs [54,55,58]. Out of the three remaining studies, one prescribed a high energy diet without supplementation [34], one study utilised 3–4 weeks of enteral nutrition follow by oral supplementation [45], and one study utilised short-term parenteral nutrition in combination with a high protein and energy diet [51].

One exercise-only intervention met the eligibility criteria. This RCT involved supervised aerobic and resistance exercise sessions of moderate intensity for 60 min three times weekly, versus a relaxation program for the control group of the same frequency and duration [41].

Across all studies, ten different methodologies were used to measure body composition (see Table 2). Most combined diet/exercise interventions used guideline-recommended measures: CT plus DXA and BIA [38], CT plus BIA [46], MRI plus BIA [29], ultrasound [37,41], or BIA alone [28,30,50]. Two diet-only studies used CT [54] or BIA [53], while the exercise-only intervention utilised DXA [51]. Anthropometric measures on their own were used predominantly in diet only studies, with the most frequent variables measured being MAMC in 12 (44%) and TSF in 11 (41%) studies. Some other anthropometric measures including calf and thigh circumference were utilised alongside guideline-recommended measures.

3.2. Quality Assessment

Plots summarising the risk of bias are presented in Figures 2 and 3. For the 19 RCTs, high risk of bias was most prevalent in domain 4 (bias in the measurement of the outcome), where assessors were often not blinded to the intervention. Almost all studies were low risk for domain 1 (randomisation and concealment processes). For domain 2, evaluating if participants and/or interventionists were blinded to the intervention allocation, the majority were allocated 'some concerns'. For the eight non-randomised studies, high risk of bias was most common in domain 3 (classification of interventions), because five of these studies were uncontrolled with no group allocations.

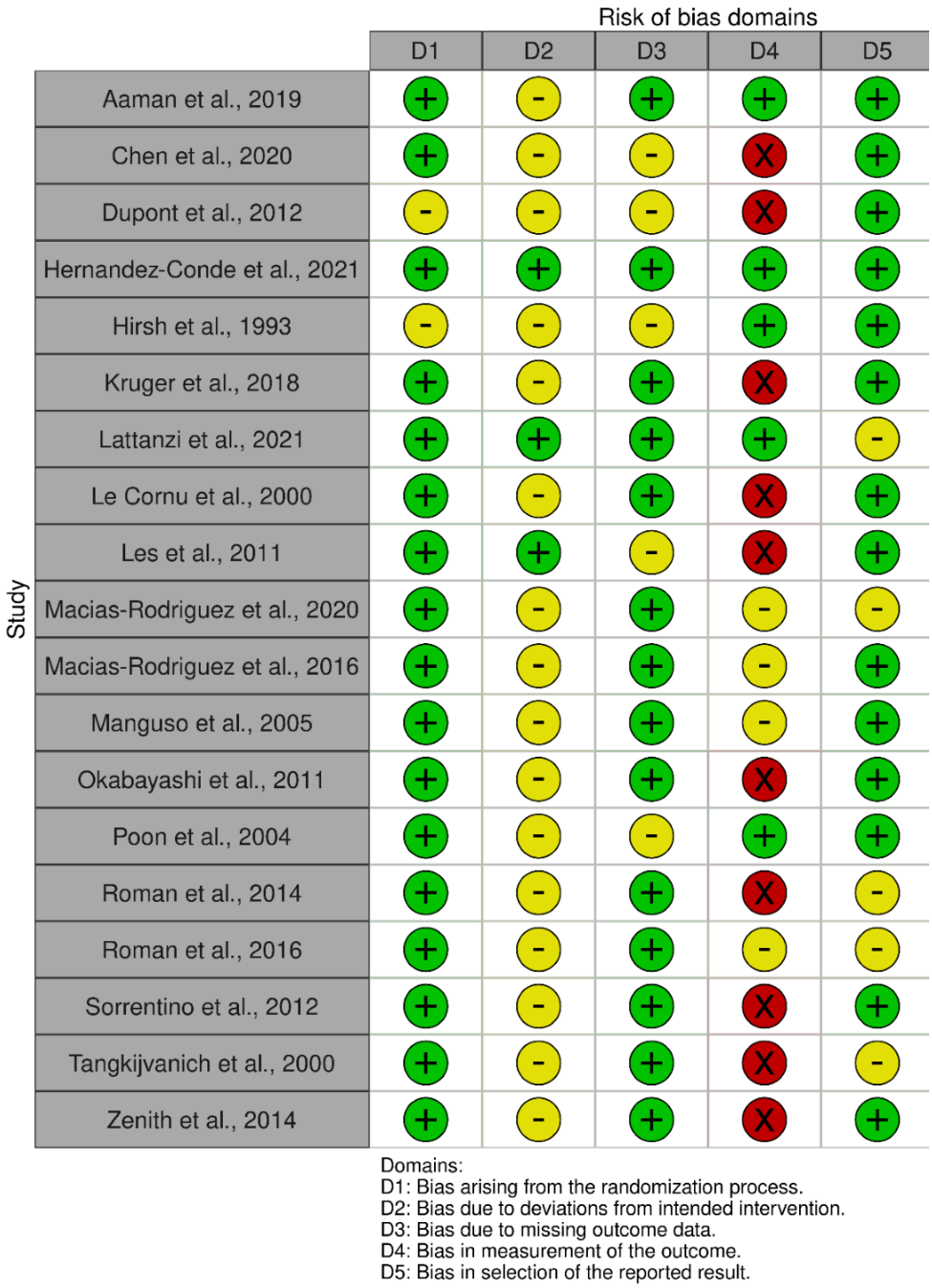


Figure 2. Risk of bias summaries for RCTs using Cochrane Risk of Bias 2 Tool [33–41,44,45,47,48,50,51,54–57].

3.3. Outcomes for Combined Diet and Exercise Intervention Studies

From the nine combined diet and exercise RCTs, four showed significant improvements in lean mass measured by CT [39], MRI [40], BIA [36], and quadricep ultrasound [35] compared to controls. One study also observed significant reductions in fat mass [35]. Three of these four studies had similar interventions of supervised, moderate intensity exercise (aerobic and/or resistance) on 3 days/week over 8–14 weeks plus targeted protein intakes above 1.2 g/kg/day through either provision of oral nutrition supplements in addition to diet, or dietetic counselling [35,37,40]. The intervention of the fourth diet and exercise RCT [39] that demonstrated an increase in skeletal muscle mass relied on frequent meals plus BCAA supplementation, with the exercise component being an increase in the number of daily steps. The participants in the control arm of this study were exposed to the

same diet and exercise intervention, but received a placebo instead of BCAAs, implicating these in the improvement in muscle mass.

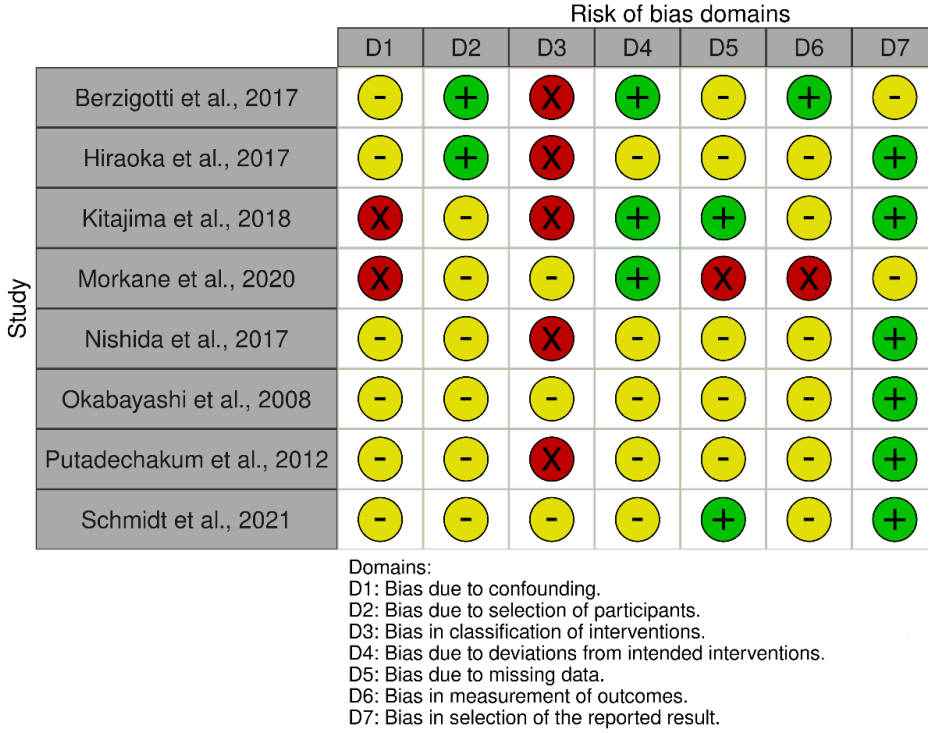


Figure 3. Risk of bias summaries for non-RCTs using Cochrane ROBINS-I (risk of bias tool to assess non-randomised studies of interventions) [32,42,43,52,53,56,58,59].

A combined diet/exercise RCT [44] which used counselling for self-directed exercise and a high protein diet with BCAA supplementation demonstrated a significant improvement compared to a diet-only control group in psoas muscle index via CT, but not in any other measures of muscle/lean mass (CT, MRI, or DXA). While the intervention group significantly increased daily number of steps compared to the control group, this small study population (n = 17) may have limited the power to detect change in some measures. This cohort of transplant candidates also had more advanced liver disease compared to the other combined interventions.

Four of the remaining diet/exercise RCTs reported significant increases in muscle mass, however this was only reported within study groups. These four studies used either supplements in combination with a diet and exercise intervention, (including HMB [38], non-alcoholic beer [37], or the amino acid leucine—a BCAA [33]); or provided dietetic counselling adjusted for BMI categories [47]. While two studies indicated good adherence to the diet and exercise interventions [37,38], the other two did not report adherence [33,47].

Both of the non-randomised combined diet/exercise intervention studies found no significant changes in lean or fat mass measured via BIA [42] or MAMC [43]. Both these studies had study population numbers of <40. One had only nine participants complete the intervention (39% attrition rate), so sample sizes may have been too small to identify significant changes [43]. The three single-arm combined diet/exercise studies assessed outpatients with compensated cirrhosis showed mixed results [32,52,53]. An intervention targeting overweight participants (BMI > 26 kg/m^2) with a 16-week program of supervised exercised with reduced caloric intake observed a significant reduction in fat mass with no significant change in lean mass [32]. A second single-arm combined intervention targeting increased step activity and BCAA supplementation reported a significant increase in muscle volume via BIA, expressed only as change ratio [52]. The final single-arm study reported no significant changes in skeletal muscle via CT, after 12 months of BCAA supplements

and prescribed bench step activity [53]. Compliance was not reported, and this study was limited by a small (n = 6) all-female cohort.

3.4. Outcomes for Diet-Only Intervention Studies

Three of nine diet-only RCTs found a significant increases in lean mass assessed by MAMC [34,48,56], while another showed a decline in the control group without change in the intervention cohort [51]. Okabayashi et al. [56] demonstrated an increase in MAMC in an intervention group using a carbohydrate enriched BCAA supplement over 12 months, combined with dietary advice to reduce energy intakes to offset the extra energy supplied with the supplement. The aim was to match dietary energy intakes to the control participants who received no supplementation; however, no dietary compliance data were reported. The second RCT [34] observed an increase in MAMC with a 12-week prescribed high energy (35–40 kcal/kg/day), low sodium diet compared to usual diet. This was a two-period cross over trial where two groups followed a prescribed diet and usual diet. The within-group change data indicated both groups significantly increased MAMC after the prescribed diet, while MAMC either declined or remained stable with usual diet. Compliance to the prescribed diet was reportedly high in both groups. The third study was of hospitalised patients. This study utilised BCAAs versus a maltodextrin supplement in the control group [48]. Short-term enteral nutrition was also provided in both groups if hepatic encephalopathy occurred and continued until oral intake was well established. The BCAA group had a significant within-group increase in MAMC, but not a significant change compared to controls.

The five remaining RCTs of diet-only interventions mostly assessed lean mass using MAMC and found no significant changes with the intervention [45,50,54,55,57]. An RCT [45] of inpatients receiving enteral nutrition for 4 weeks as a component of their intervention found no significant changes in MAMC or TSF compared to inpatients receiving a usual hospital diet. A four week diet-only RCT [57] provided an isocaloric diet for both intervention and controls (2000 kcal and 80 g of protein/day), with the intervention group receiving BCAA supplementation and a reduced diet to achieve the same energy and protein intake as the control. Only within-group changes were reported and no diet compliance data were presented. MAMC did not significantly change in participants given the BCAA supplement over 12 months compared to usual diet [50]. Average intakes declined marginally in both groups even though BCAA compliance was satisfactory. A study in transplant candidates [55] with a MAMC below the 25th percentile also saw no improvement in this measure following supplementation and dietary counselling until transplantation, versus dietary counselling alone. The RCT by Hirsch et al. [54]. provided supplements over 12 months to patients with decompensated cirrhosis. While mean oral intakes appeared significantly higher in the intervention versus control, there were no significant changes in MAMC or TSF.

The final diet-only RCT [51] evaluated the effect of three diets in people with decompensated cirrhosis and ascites on lean and fat mass assessed by anthropometry. The first diet (Group A) prescribed 24 h of parenteral nutrition in addition to a high energy and protein diet with monthly dietitian advice. Group B received the same diet without parenteral nutrition while the third (control) group were prescribed a "sodium free" diet with dietitian advice. The control group had a significant decline in TSF and MAC compared to Groups A and B. MAMC was not reported in this study. Unfortunately, the control group had mean dietary protein and energy (0.6 ± 3 g/kg/day and 25 ± 8 kcal/kg/day respectively), considerably below guidelines for decompensated cirrhosis [16]. This is likely the cause for the changes in the control group and highlights the potential negative impact of a restrictive low sodium diet without a protein or energy prescription in patients with decompensated cirrhosis.

Three non-randomised or single arm studies of diet-only interventions yielded mixed results [49,58,59]. One non-randomised study [49] assessed patients who had undergone surgery for HCC and reported a within-group increase in MAMC after 6 months of BCAA

supplementation twice daily, using a historical control group who received no supplementation. No MAMC data were reported for the control and MAMC was not compared between groups. One of the two single-arm diet-only interventions [58] saw a significant improvement in lean mass via BIA with a soy-based nutritional supplement over 8 weeks plus usual diet. Dietary intake changes were not reported. Participants had predominantly Child-Pugh A cirrhosis, allowing a reasonably reliable interpretation of BIA. The final single-arm diet-only study [59] reported no significant change in skeletal muscle via CT. The intervention of BCAA supplementation over 48-weeks was said to have 100% adherence to the supplement. Intramuscular adipose tissue was also assessed via CT with no significant change observed.

3.5. Outcomes for Exercise-Only Interventions

The one exercise-only study [41] was an RCT involving 12 weeks of supervised moderate intensity aerobic exercise 3 days/week, compared to a "sham intervention" of relaxation exercises. The exercise group, which reported high attendance rates, significantly increased lean mass via DXA. Additionally, there was a significant within-group reduction in fat mass in the intervention group as well as an increase in upper thigh circumference and reduction in mid-arm circumference. There were no significant changes in the "sham" group, but comparisons between the active and sham groups were not reported.

4. Discussion

The aim of this systematic review was to assess the impact of diet and/or exercise interventions on body composition in patients with liver cirrhosis. While published reviews exist on nutrition and/or exercise interventions in cirrhosis, there are none, to our knowledge, that reviewed studies which specifically measured body composition across both diet and exercise interventions. Secondly, this review also sought to determine the effect of these interventions in patients with cirrhosis and obesity, given the increasing prevalence of obesity and the risks associated with this [12], versus the potential deleterious impact of calorie restriction on muscle mass in this population.

Unfortunately, the 27 studies identified for this review were too heterogeneous in terms of design and outcome measures to allow meta-analysis. Small study size and failure to report adherence to interventions also impacted data synthesis. Nonetheless, on systematic review, the combined diet and exercise interventions appeared to show the greatest potential to increase muscle mass. To demonstrate an increase in muscle mass with exercise, these interventions needed to be of $\geq$8 weeks duration and comprise 30–60 min of moderate intensity supervised exercise (aerobic and/or resistance), on at least 3 days per week combined with protein intakes of 1.2–2 g/kg/day. In addition, there appeared to be a benefit to muscle mass from BCAA supplementation [35,36,39,40]. Interestingly, several of the combined RCTs [35,37–39,44] provided the control group with either a diet or exercise intervention. Based on these studies it appears that there is a synergistic effect when both diet and exercise interventions are delivered to increase muscle mass.

Obesity is known to impact patients with cirrhosis as an important contributor to progression of liver disease. In patients undergoing liver transplant, severe obesity (BMI > 35 kg/m^2) increases the risk of peri-transplant complications and death [61]. The prevalence of obesity is increasing in the whole population and in patients with advanced liver disease, and so the impacts of obesity in patients with advanced liver disease are likely to become increasingly important. A challenge addressing obesity in patients with cirrhosis is that the catabolic metabolism found in advanced liver disease could potentially result in significant muscle loss with calorie restriction. Of the 27 studies included in this systematic review, 19 reported on dry weight BMI, with the mean BMI of patients in 13 of these studies being in the overweight [33–40] or obese [32,41–44] ranges. Only two studies reported on changes in fat mass [32,39]. In the RCT by Hernandez-Conde and colleagues [39], the mean BMI of patients in intervention and control arm were in the overweight range. Although weight loss was not a specific goal of their study, they showed that a combined intervention

of diet, exercise, and BCAA supplementation led to a reduction in fat mass while muscle mass improved. The one study targeting weight loss in overweight and obese patients with cirrhosis was promising in that it demonstrated a fall in body weight with maintenance of lean mass after 16-weeks of a combined intervention of exercise with a reduced energy, moderate protein diet [32].

In relation to the heterogeneity of studies, one issue that impacted the ability to synthesise the findings of this review was that 10 different methods of assessing body composition were used across the 27 studies. Current guidelines recommend CT or MRI as optimal body composition assessment methods, in part because they are less impacted by the fluid overload and ascites that occur in decompensated cirrhosis than some of the other methods [16,19]. While CT/MRI are expensive and not always available, they are often part of standard of care for patients undergoing transplant evaluation to assess hepatic vasculature or for HCC monitoring. Although these routine measures are not performed specifically to assess body composition, they can additionally be used to assess muscle and fat mass; and it is possible for allied health clinicians to perform these analyses [62].

When abdominal CT or MRI are not available, guidelines recommend using DXA and BIA to assess body composition, on the provision that fluid retention is not an issue [19]; however, this restricts their utility in the group of cirrhotic patients most at risk of sarcopenia, those with decompensated disease. Muscle mass quantification by DXA has been shown to correlate with CT in cirrhosis [63]. Ultrasound is promising yet requires further exploration in this population [6,16]. While the accuracy of BIA can be affected by hydration [21], the use of Phase Angle from BIA may provide a more reliable assessment of nutritional status in cirrhosis than other BIA modalities [19], with results comparable to CT [64]. Several studies only utilised MAMC and TSF as outcome measures, particularly in the diet only interventions. Anthropometry is routinely used in the clinical assessment of nutritional status. However, the utility in clinical studies is less clear as these measures suffer in regard to reliability [18], and cannot distinguish small changes in body composition [7,26]. This makes them less than ideal for studies conducted over 8–12 weeks like a number of the studies reported here. Additional issues with their use in the studies included in this review were that outcome assessors were frequently not blinded to intervention arm, or for these very operator dependent measures, that several assessors may have been involved in the serial measurements. This increases the impact of interobserver variability on findings. Interestingly, while there is a strong body of evidence indicating the deleterious effects of sarcopenia in cirrhosis, very few of the included studies assessed if the patient's level of baseline muscle mass was indicative of sarcopenia prior to conducting the intervention. This highlights the need for future studies to evaluate baseline muscle mass and therefore sarcopenia to understand the true effect of diet and/or exercise interventions.

An additional issue was that most diet and exercise studies in cirrhosis have small study populations with body composition measures generally underpowered and as mentioned, were often included as secondary outcomes of the studies. Some of the challenges to increasing participant numbers in studies in this area are the complexities of conducting lifestyle interventions in a population with advanced liver disease who may be quite unwell. Another factor which can impact drop-out rates in this population is inclusion of participants who are potential transplant candidates. In one of the RCTs included in this review [43], a 6-week exercise intervention was completed in just over half (56%) of potential liver transplant recipients, and this was largely because of study participants receiving a liver transplant rather than not adhering to the program. We faced a similar issue in an 8-week pilot feasibility RCT of exercise in patients on a liver transplant waiting list [65], only 50% of participants completed the study, largely because participants received their liver transplant within the study period.

This review also highlighted the sparsity of relevant intervention studies which have targeted patients with decompensated liver disease. Patients with decompensated disease are a complex and high-risk population, who are more likely to experience muscle wasting

and adverse outcomes [6]. Chen et al. [44], is one of the few studies in this review that included only decompensated cirrhosis patients that used a combined diet and exercise intervention. This study was small; however, they were able to demonstrate that home-based exercise is safe in this population. This is promising, and future studies should focus on these populations to better understand how body composition can be improved pre-transplant to improve morbidity and mortality.

While this systematic literature review focused on changes in body composition measured using methodology validated in liver cirrhosis, there are other diet and or exercise intervention studies that have added value to the management of patients with advanced liver disease. Several exercise RCTs in patients with cirrhosis have measured aspects of physical performance including strength, exercise capacity, and/or physical function and therefore did not meet the inclusion criteria for this review. Measures such as hand grip strength, anaerobic threshold by cardiopulmonary exercise testing and functional performance assessments such as the Short Physical Performance Battery and the Liver Frailty Index have demonstrated associations with patient outcomes [5,66–68]. They can be useful as screening tools to identify patients at risk of complications [69] and are recommended as part of the evaluation of nutritional status in people with cirrhosis [16,19,26]. Consideration should be given to including these measures in future studies that address body composition alongside functional status.

The field of diet and exercise interventions in patients with cirrhosis is obviously at an early and evolving stage. An important goal for future studies is to determine the significance of modest improvements in body composition both in terms of clinical outcomes, but also in patient-important outcomes and their quality of life. Given the potential range and combination of diet and exercise interventions, defining minimal clinically important differences for muscle and fat mass and thresholds for adverse outcomes patients should be a goal to facilitate comparisons between interventions.

5. Conclusions

In summary, effective interventions to improve body composition in cirrhosis appear more likely to succeed if diet and exercise components are combined. There remains a paucity of studies in patients with cirrhosis and obesity despite the increasing prevalence of obesity in this population. At present, the evidence supporting diet and exercise approaches to improve body composition in cirrhosis is impacted by underpowered, short-term interventions. Future research should be directed at appropriately powered combined diet and exercise RCTs of at least 8 weeks duration. Ideally assessments of changes in muscle mass, particularly in patients with decompensated cirrhosis should rely on guideline-recommended methods in this population, specifically CT or MRI. These studies should ideally be large enough to allow for the potentially high rates of patient drop-out and include formal assessments of patient adherence to interventions to identify strategies that do and do not work in this cohort. An important goal for future studies should be to determine what are clinically meaningful changes in body composition in patients with cirrhosis as this will facilitate comparison between intervention strategies. These approaches will help clarify if sarcopenia and sarcopenic obesity are modifiable risk factors in cirrhosis.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/nu14163365/s1, Supplementary File S1: PRISMA 2020 Checklist for Systematic Reviews; Supplementary File S2: Systematic review search strategy.

Author Contributions: Conceptualisation, H.E.J., T.G.T., I.J.H., and H.L.M.; methodology, H.E.J., T.G.T., J.T.K., I.J.H., and H.L.M.; software, H.E.J.; validation, H.E.J., G.A.M., I.J.H., and H.L.M.; formal analysis, H.E.J., T.G.T., I.J.H., and H.L.M.; investigation, H.E.J., T.G.T., and H.L.M.; resources, H.E.J., I.J.H., and H.L.M.; data curation, H.E.J., T.G.T., and H.L.M.; writing—original draft preparation, H.E.J., I.J.H., and H.L.M.; writing—review and editing, H.E.J., T.G.T., S.E.K., G.A.M., J.S.C., J.T.K., I.J.H., and H.L.M.; visualisation, H.E.J., I.J.H., G.A.M., and H.L.M.; supervision, I.J.H., G.A.M., and H.L.M.; project administration, H.E.J. and H.L.M.; funding acquisition, H.E.J. All authors have read and agreed to the published version of the manuscript.

Funding: Heidi Johnston was supported by an Australian Government Research Training Program and Living Stipend Scholarship via the University of Queensland.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: We thank the University of Queensland Librarian Marcos Riba for assistance refining the search strategy process and the Department of Nutrition and Dietetics at the Princess Alexandra Hospital for in-kind support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moon, A.M.; Singal, A.G.; Tapper, E.B. Contemporary epidemiology of chronic liver disease and cirrhosis. *Clin. Gastroenterol. Hepatol.* **2020**, *18*, 2650–2666. [CrossRef] [PubMed]
2. D'Amico, G.; Garcia-Tsao, G.; Pagliaro, L. Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies. *J. Hepatol.* **2006**, *44*, 217–231. [CrossRef]
3. Cheung, K.; Lee, S.S.; Raman, M. Prevalence and mechanisms of malnutrition in patients with advanced liver disease, and nutrition management strategies. *Clin. Gastroenterol. Hepatol.* **2012**, *10*, 117–125. [CrossRef] [PubMed]
4. Bhanji, R.A.; Carey, E.J.; Yang, L.; Watt, K.D. The long winding road to transplant: How sarcopenia and debility impact morbidity and mortality on the waitlist. *Clin. Gastroenterol. Hepatol.* **2017**, *15*, 1492–1497. [CrossRef] [PubMed]
5. Sinclair, M.; Poltavskiy, E.; Dodge, J.L.; Lai, J.C. Frailty is independently associated with increased hospitalisation days in patients on the liver transplant waitlist. *World J. Gastroenterol.* **2017**, *23*, 899. [CrossRef]
6. Tandon, P.; Montano-Loza, A.J.; Lai, J.C.; Dasarathy, S.; Merli, M. Sarcopenia and frailty in decompensated cirrhosis. *J. Hepatol.* **2021**, *75*, S147–S162. [CrossRef]
7. Cruz-Jentoft, A.J.; Baeyens, J.P.; Bauer, J.M.; Boirie, Y.; Cederholm, T.; Landi, F.; Martin, F.C.; Michel, J.-P.; Rolland, Y.; Schneider, S.M.; et al. Sarcopenia: European consensus on definition and diagnosisReport of the European Working Group on Sarcopenia in Older People. *Age Ageing* **2010**, *39*, 412–423. [CrossRef]
8. Laube, R.; Wang, H.; Park, L.; Heyman, J.K.; Vidot, H.; Majumdar, A.; Strasser, S.I.; McCaughan, G.W.; Liu, K. Frailty in advanced liver disease. *Liver Int.* **2018**, *38*, 2117–2128. [CrossRef]
9. Kim, G.; Kang, S.H.; Kim, M.Y.; Baik, S.K. Prognostic value of sarcopenia in patients with liver cirrhosis: A systematic review and meta-analysis. *PLoS ONE* **2017**, *12*, e0186990. [CrossRef]
10. Van Vugt, J.; Levolger, S.; de Bruin, R.; van Rosmalen, J.; Metselaar, H.; IJzermans, J. Systematic review and meta-analysis of the impact of computed tomography–assessed skeletal muscle mass on outcome in patients awaiting or undergoing liver transplantation. *Am. J. Transplant.* **2016**, *16*, 2277–2292. [CrossRef]
11. Berzigotti, A.; Garcia-Tsao, G.; Bosch, J.; Grace, N.D.; Burroughs, A.K.; Morillas, R.; Escorsell, A.; Garcia-Pagan, J.C.; Patch, D.; Matloff, D.S. Obesity is an independent risk factor for clinical decompensation in patients with cirrhosis. *Hepatology* **2011**, *54*, 555–561. [CrossRef] [PubMed]
12. Montano-Loza, A.J.; Angulo, P.; Meza-Junco, J.; Prado, C.M.; Sawyer, M.B.; Beaumont, C.; Esfandiari, N.; Ma, M.; Baracos, V.E. Sarcopenic obesity and myosteatosis are associated with higher mortality in patients with cirrhosis. *J. Cachexia Sarcopenia Muscle* **2016**, *7*, 126–135. [CrossRef] [PubMed]
13. Spengler, E.K.; O'Leary, J.G.; Te, H.S.; Rogal, S.; Pillai, A.A.; Al-Osaimi, A.; Desai, A.; Fleming, J.N.; Ganger, D.; Seetharam, A.; et al. Liver Transplantation in the Obese Cirrhotic Patient. *Transplantation* **2017**, *101*, 2288–2296. [CrossRef] [PubMed]
14. Vidot, H.; Kline, K.; Cheng, R.; Finegan, L.; Lin, A.; Kempler, E.; Strasser, S.I.; Bowen, D.G.; McCaughan, G.W.; Carey, S. The relationship of obesity, nutritional status and muscle wasting in patients assessed for liver transplantation. *Nutrients* **2019**, *11*, 2097. [CrossRef]
15. Calzadilla-Bertot, L.; Jeffrey, G.P.; Jacques, B.; McCaughan, G.; Crawford, M.; Angus, P.; Jones, R.; Gane, E.; Munn, S.; Macdonald, G. Increasing incidence of nonalcoholic steatohepatitis as an indication for liver transplantation in Australia and New Zealand. *Liver Transpl.* **2019**, *25*, 25–34. [CrossRef]
16. Plauth, M.; Bernal, W.; Dasarathy, S.; Merli, M.; Plank, L.D.; Schütz, T.; Bischoff, S.C. European Society of Enteral and Parenteral Nutrition Guideline on Clinical Nutrition in Liver Disease. *Clin. Nutr.* **2019**, *38*, 485–521. [CrossRef]
17. Morgan, M.Y.; Madden, A.M.; Soulsby, C.T.; Morris, R.W. Derivation and validation of a new global method for assessing nutritional status in patients with cirrhosis. *Hepatology* **2006**, *44*, 823–835. [CrossRef]
18. Ulijaszek, S.J.; Kerr, D.A. Anthropometric measurement error and the assessment of nutritional status. *Br. J. Nutr.* **1999**, *82*, 165–177. [CrossRef]
19. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on nutrition in chronic liver disease. *J. Hepatol.* **2019**, *70*, 172–193. [CrossRef]

20. Sinclair, M.; Hoermann, R.; Peterson, A.; Testro, A.; Angus, P.W.; Hey, P.; Chapman, B.; Gow, P.J. Use of dual X-ray absorptiometry in men with advanced cirrhosis to predict sarcopenia-associated mortality risk. *Liver Int.* **2019**, *39*, 1089–1097. [CrossRef]
21. Morgan, M.Y.; Madden, A.M.; Jennings, G.; Elia, M.; Fuller, N.J. Two-component models are of limited value for the assessment of body composition in patients with cirrhosis. *Am. J. Clin. Nutr.* **2006**, *84*, 1151–1162. [CrossRef] [PubMed]
22. Bowen, T.S.; Schuler, G.; Adams, V. Skeletal muscle wasting in cachexia and sarcopenia: Molecular pathophysiology and impact of exercise training. *J. Cachexia Sarcopenia Muscle* **2015**, *6*, 197–207. [CrossRef] [PubMed]
23. Williams, F.R.; Berzigotti, A.; Lord, J.M.; Lai, J.C.; Armstrong, M.J. Impact of exercise on physical frailty in patients with chronic liver disease. *Aliment. Pharmacol. Ther.* **2019**, *50*, 988–1000. [CrossRef] [PubMed]
24. Toshikuni, N.; Arisawa, T.; Tsutsumi, M. Nutrition and exercise in the management of liver cirrhosis. *World J. Gastroenterol.* **2014**, *20*, 7286–7297. [CrossRef] [PubMed]
25. Ooi, P.H.; Gilmour, S.M.; Yap, J.; Mager, D.R. Effects of branched chain amino acid supplementation on patient care outcomes in adults and children with liver cirrhosis: A systematic review. *Clin. Nutr. ESPEN* **2018**, *28*, 41–51. [CrossRef]
26. Lai, J.C.; Tandon, P.; Bernal, W.; Tapper, E.B.; Ekong, U.; Dasarathy, S.; Carey, E.J. Malnutrition, Frailty, and Sarcopenia in Patients With Cirrhosis: 2021 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology* **2021**, *74*, 1611–1644. [CrossRef] [PubMed]
27. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef]
28. The EndNote Team. *EndNote*; Endnote X9; Clarivate: Philadelphia, PA, USA, 2013.
29. Ouzzani, M.; Hammady, H.; Fedorowicz, Z.; Elmagarmid, A. Rayyan—A web and mobile app for systematic reviews. *Syst. Rev.* **2016**, *5*, 210. [CrossRef]
30. Sterne, J.A.; Savović, J.; Page, M.J.; Elbers, R.G.; Blencowe, N.S.; Boutron, I.; Cates, C.J.; Cheng, H.-Y.; Corbett, M.S.; Eldridge, S.M. RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **2019**, *366*, l4898. [CrossRef]
31. Sterne, J.A.; Hernán, M.A.; McAleenan, A.; Reeves, B.C.; Higgins, J.P. ROBINS-I: A tool for assessing risk of bias in non-randomized studies of interventions. *BMJ.* **2016**, *355*, i4919. [CrossRef]
32. Berzigotti, A.; Albillos, A.; Villanueva, C.; Genescá, J.; Ardevol, A.; Augustín, S.; Calleja, J.L.; Bañares, R.; García-Pagán, J.C.; Mesonero, F. Effects of an intensive lifestyle intervention program on portal hypertension in patients with cirrhosis and obesity: The SportDiet study. *Hepatology* **2017**, *65*, 1293–1305. [CrossRef] [PubMed]
33. Román, E.; Torrades, M.T.; Nadal, M.J.; Cárdenas, G.; Nieto, J.C.; Vidal, S.; Bascunana, H.; Juárez, C.; Guarner, C.; Córdoba, J. Randomized pilot study: Effects of an exercise programme and leucine supplementation in patients with cirrhosis. *Dig. Dis. Sci.* **2014**, *59*, 1966–1975. [CrossRef] [PubMed]
34. Manguso, F.; D'ambra, G.; Menchise, A.; Sollazzo, R.; D'agostino, L. Effects of an appropriate oral diet on the nutritional status of patients with HCV-related liver cirrhosis: A prospective study. *Clin. Nutr.* **2005**, *24*, 751–759. [CrossRef] [PubMed]
35. Zenith, L.; Meena, N.; Ramadi, A.; Yavari, M.; Harvey, A.; Carbonneau, M.; Ma, M.; Abraldes, J.G.; Paterson, I.; Haykowsky, M.J. Eight weeks of exercise training increases aerobic capacity and muscle mass and reduces fatigue in patients with cirrhosis. *Clin. Gastroenterol. Hepatol.* **2014**, *12*, 1920–1926.e2. [CrossRef]
36. Macías-Rodríguez, R.U.; Ilarraza-Lomelí, H.; Ruiz-Margáin, A.; Ponce-de-León-Rosales, S.; Vargas-Vorácková, F.; García-Flores, O.; Torre, A.; Duarte-Rojo, A. Changes in hepatic venous pressure gradient induced by physical exercise in cirrhosis: Results of a pilot randomized open clinical trial. *Clin. Transl. Gastroenterol.* **2016**, *7*, e180. [CrossRef]
37. Macías-Rodríguez, R.U.; Ruiz-Margáin, A.; Román-Calleja, B.M.; Espin-Nasser, M.E.; Flores-García, N.C.; Torre, A.; Galicia-Hernández, G.; Rios-Torres, S.L.; Fernández-del-Rivero, G.; Orea-Tejeda, A. Effect of non-alcoholic beer, diet and exercise on endothelial function, nutrition and quality of life in patients with cirrhosis. *World J. Hepatol.* **2020**, *12*, 1299. [CrossRef]
38. Lattanzi, B.; Bruni, A.; Di Cola, S.; Molfino, A.; De Santis, A.; Muscaritoli, M.; Merli, M. The Effects of 12-Week Beta-Hydroxy-Beta-Methylbutyrate Supplementation in Patients with Liver Cirrhosis: Results from a Randomized Controlled Single-Blind Pilot Study. *Nutrients* **2021**, *13*, 2296. [CrossRef]
39. Hernández-Conde, M.; Llop, E.; Gómez-Pimpollo, L.; Carrillo, C.F.; Rodríguez, L.; Van Den Brule, E.; Perelló, C.; López-Gómez, M.; Abad, J.; Martínez-Porras, J.L. Adding Branched-Chain Amino Acids to an Enhanced Standard-of-Care Treatment Improves Muscle Mass of Cirrhotic Patients With Sarcopenia: A Placebo-Controlled Trial. *Off. J. Am. Coll. Gastroenterol.* **2021**, *116*, 2241–2249. [CrossRef]
40. Aamann, L.; Dam, G.; Borre, M.; Drljevic-Nielsen, A.; Overgaard, K.; Andersen, H.; Vilstrup, H.; Aagaard, N.K. Resistance training increases muscle strength and muscle size in patients with liver cirrhosis. *Clin. Gastroenterol. Hepatol.* **2019**, *18*, 1179–1187. [CrossRef]
41. Román, E.; García-Galcerán, C.; Torrades, T.; Herrera, S.; Marín, A.; Doñate, M.; Alvarado-Tapias, E.; Malouf, J.; Nácher, L.; Serra-Grima, R. Effects of an exercise programme on functional capacity, body composition and risk of falls in patients with cirrhosis: A randomized clinical trial. *PLoS ONE* **2016**, *11*, e0151652. [CrossRef]
42. Schmidt, N.P.; Fernandes, S.A.; Silveira, A.T.; Rayn, R.G.; Henz, A.C.; Rossi, D.; Galant, L.H.; Marroni, C.A. Nutritional and functional rehabilitation in cirrhotic patients. *J. Gastroenterol. Hepatol. Res.* **2021**, *10*, 3470–3477.
43. Morkane, C.M.; Kearney, O.; Bruce, D.A.; Melikian, C.N.; Martin, D.S. An outpatient hospital-based exercise training program for patients with cirrhotic liver disease awaiting transplantation: A feasibility trial. *Transplantation* **2020**, *104*, 97–103. [CrossRef] [PubMed]

44. Chen, H.W.; Ferrando, A.; White, M.G.; Dennis, R.A.; Xie, J.; Pauly, M.; Park, S.; Bartter, T.; Dunn, M.A.; Ruiz-Margain, A. Home-Based Physical Activity and Diet Intervention to Improve Physical Function in Advanced Liver Disease: A Randomized Pilot Trial. *Dig. Dis. Sci.* **2020**, *65*, 3350–3359. [CrossRef] [PubMed]
45. Dupont, B.; Dao, T.; Joubert, C.; Dupont-Lucas, C.; Gloro, R.; Nguyen-Khac, E.; Beaujard, E.; Mathurin, P.; Vastel, E.; Musikas, M. Randomised clinical trial: Enteral nutrition does not improve the long-term outcome of alcoholic cirrhotic patients with jaundice. *Aliment. Pharmacol. Ther.* **2012**, *35*, 1166–1174. [CrossRef]
46. Debette-Gratien, M.; Tabouret, T.; Antonini, M.-T.; Dalmay, F.; Carrier, P.; Legros, R.; Jacques, J.; Vincent, F.; Sautereau, D.; Samuel, D. Personalized adapted physical activity before liver transplantation: Acceptability and results. *Transplantation* **2015**, *99*, 145–150. [CrossRef]
47. Kruger, C.; McNeely, M.L.; Bailey, R.J.; Yavari, M.; Abraldes, J.G.; Carbonneau, M.; Newnham, K.; DenHeyer, V.; Ma, M.; Thompson, R. Home exercise training improves exercise capacity in cirrhosis patients: Role of exercise adherence. *Sci. Rep.* **2018**, *8*, 99. [CrossRef]
48. Les, I.; Doval, E.; García-Martínez, R.; Planas, M.; Cárdenas, G.; Gómez, P.; Flavià, M.; Jacas, C.; Mínguez, B.; Vergara, M. Effects of branched-chain amino acids supplementation in patients with cirrhosis and a previous episode of hepatic encephalopathy: A randomized study. *Off. J. Am. Coll. Gastroenterol.* **2011**, *106*, 1081–1088. [CrossRef]
49. Okabayashi, T.; Nishimori, I.; Sugimoto, T.; Iwasaki, S.; Akisawa, N.; Maeda, H.; Ito, S.; Onishi, S.; Ogawa, Y.; Kobayashi, M. The benefit of the supplementation of perioperative branched-chain amino acids in patients with surgical management for hepatocellular carcinoma: A preliminary study. *Dig. Dis. Sci.* **2008**, *53*, 204–209. [CrossRef]
50. Poon, R.P.; Yu, W.C.; Fan, S.T.; Wong, J. Long-term oral branched chain amino acids in patients undergoing chemoembolization for hepatocellular carcinoma: A randomized trial. *Aliment. Pharmacol. Ther.* **2004**, *19*, 779–788. [CrossRef]
51. Sorrentino, P.; Castaldo, G.; Tarantino, L.; Bracigliano, A.; Perrella, A.; Perrella, O.; Fiorentino, F.; Vecchione, R.; D'Angelo, S. Preservation of nutritional-status in patients with refractory ascites due to hepatic cirrhosis who are undergoing repeated paracentesis. *J. Gastroenterol. Hepatol.* **2012**, *27*, 813–822. [CrossRef]
52. Hiraoka, A.; Michitaka, K.; Kiguchi, D.; Izumoto, H.; Ueki, H.; Kaneto, M.; Kitahata, S.; Aibiki, T.; Okudaira, T.; Tomida, H. Efficacy of branched-chain amino acid supplementation and walking exercise for preventing sarcopenia in patients with liver cirrhosis. *Eur. J. Gastroenterol. Hepatol.* **2017**, *29*, 1416–1423. [CrossRef] [PubMed]
53. Nishida, Y.; Ide, Y.; Okada, M.; Otsuka, T.; Eguchi, Y.; Ozaki, I.; Tanaka, K.; Mizuta, T. Effects of home-based exercise and branched-chain amino acid supplementation on aerobic capacity and glycemic control in patients with cirrhosis. *Hepatol. Res.* **2017**, *47*, E193–E200. [CrossRef] [PubMed]
54. Hirsch, S.; Bunout, D.; De La Maza, P.; Iturriaga, H.; Petermann, M.; Icazar, G.; Gattas, V.; Ugarte, G. Controlled trial on nutrition supplementation in outpatients with symptomatic alcoholic cirrhosis. *J. Parenter. Enter. Nutr.* **1993**, *17*, 119–124. [CrossRef] [PubMed]
55. Le Cornu, K.A.; McKiernan, F.J.; Kapadia, S.A.; Neuberger, J.M. A prospective randomized study of preoperative nutritional supplementation in patients awaiting elective Orthotopic liver Transplantation. *Transplantation* **2000**, *69*, 1364–1369. [CrossRef]
56. Okabayashi, T.; Iyoki, M.; Sugimoto, T.; Kobayashi, M.; Hanazaki, K. Oral supplementation with carbohydrate-and branched-chain amino acid-enriched nutrients improves postoperative quality of life in patients undergoing hepatic resection. *Amino Acids* **2011**, *40*, 1213–1220. [CrossRef]
57. Tangkijvanich, P.; Mahachai, V.; Wittayalertpanya, S.; Ariyawongsopon, V.; Isarasena, S. Short-term effects of branched-chain amino acids on liver function tests in cirrhotic patients. *Southeast Asian J. Trop. Med. Public Health* **2000**, *31*, 152–157.
58. Putadechakum, S.; Klangjareonchai, T.; Soponsaritsuk, A.; Roongpisuthipong, C. Nutritional status assessment in cirrhotic patients after protein supplementation. *Int. Sch. Res. Not.* **2012**, *2012*, 690402. [CrossRef]
59. Kitajima, Y.; Takahashi, H.; Akiyama, T.; Murayama, K.; Iwane, S.; Kuwashiro, T.; Tanaka, K.; Kawazoe, S.; Ono, N.; Eguchi, T. Supplementation with branched-chain amino acids ameliorates hypoalbuminemia, prevents sarcopenia, and reduces fat accumulation in the skeletal muscles of patients with liver cirrhosis. *J. Gastroenterol.* **2018**, *53*, 427–437. [CrossRef]
60. Pugh, R.; Murray-Lyon, I.; Dawson, J.; Pietroni, M.; Williams, R. Transection of the oesophagus for bleeding oesophageal varices. *J. Br. Surg.* **1973**, *60*, 646–649. [CrossRef]
61. Molina Raya, A.; García Navarro, A.; San Miguel Méndez, C.; Domínguez Bastante, M.; Villegas Herrera, M.T.; Granero, K.; Becerra Massare, A.; Villar Del Moral, J.M.; Expósito, M.; Fundora Suárez, Y. Influence of Obesity on Liver Transplantation Outcomes. *Transplant. Proc.* **2016**, *48*, 2503–2505. [CrossRef]
62. Johnston, H.E.; de Crom, T.; Hargrave, C.; Adhyaru, P.; Woodward, A.J.; Pang, S.; Ali, A.; Coombes, J.S.; Keating, S.E.; McLean, K. The inter-and intrarater reliability and feasibility of dietetic assessment of sarcopenia and frailty in potential liver transplant recipients: A mixed-methods study. *Clin. Transplant.* **2021**, *35*, e14185. [CrossRef]
63. Georgiou, A.; Papatheodoridis, G.V.; Alexopoulou, A.; Deutsch, M.; Vlachogiannakos, I.; Ioannidou, P.; Papageorgiou, M.-V.; Papadopoulos, N.; Yannakoulia, M.; Kontogianni, M.D. Validation of cutoffs for skeletal muscle mass index based on computed tomography analysis against dual energy X-ray absorptiometry in patients with cirrhosis: The KIRRHOS study. *Ann. Gastroenterol.* **2020**, *33*, 80. [CrossRef] [PubMed]
64. Ruiz-Margáin, A.; Xie, J.J.; Román-Calleja, B.M.; Pauly, M.; White, M.G.; Chapa-Ibargüengoitia, M.; Campos-Murguía, A.; González-Regueiro, J.A.; Macias-Rodríguez, R.U.; Duarte-Rojo, A. Phase Angle From Bioelectrical Impedance for the Assessment of Sarcopenia in Cirrhosis With or Without Ascites. *Clin. Gastroenterol. Hepatol.* **2021**, *19*, 1941–1949.e2. [CrossRef]

65. Wallen, M.P.; Keating, S.E.; Hall, A.; Hickman, I.J.; Pavey, T.G.; Woodward, A.J.; Skinner, T.L.; Macdonald, G.A.; Coombes, J.S. Exercise training is safe and feasible in patients awaiting liver transplantation: A Pilot Randomized Controlled Trial. *Liver Transpl.* **2019**, *25*, 1576–1580. [CrossRef] [PubMed]
66. Lai, J.C.; Feng, S.; Terrault, N.A.; Lizaola, B.; Hayssen, H.; Covinsky, K. Frailty predicts waitlist mortality in liver transplant candidates. *Am. J. Transplant.* **2014**, *14*, 1870–1879. [CrossRef] [PubMed]
67. Sinclair, M.; Chapman, B.; Hoermann, R.; Angus, P.W.; Testro, A.; Scodellaro, T.; Gow, P.J. Handgrip strength adds more prognostic value to the Model for End-Stage Liver Disease score than imaging-based measures of muscle mass in men with cirrhosis. *Liver Transpl.* **2019**, *25*, 1480–1487. [CrossRef] [PubMed]
68. Tandon, P.; Tangri, N.; Thomas, L.; Zenith, L.; Shaikh, T.; Carbonneau, M.; Ma, M.; Bailey, R.J.; Jayakumar, S.; Burak, K.W.; et al. A Rapid Bedside Screen to Predict Unplanned Hospitalization and Death in Outpatients With Cirrhosis: A Prospective Evaluation of the Clinical Frailty Scale. *Am. J. Gastroenterol.* **2016**, *111*, 1759–1767. [CrossRef]
69. Lai, J.C.; Dodge, J.L.; Kappus, M.R.; Dunn, M.A.; Volk, M.L.; Duarte-Rojo, A.; Ganger, D.R.; Rahimi, R.S.; McCulloch, C.E.; Haugen, C.E. Changes in frailty are associated with waitlist mortality in patients with cirrhosis. *J. Hepatol.* **2020**, *73*, 575–581. [CrossRef]

The Journal of Nursing Research ■ VOL. 32, NO. 1, FEBRUARY 2024

DOI: https://doi.org/10.1097/jnr.0000000000000595

OPEN

Effects of a Health Literacy Education Program on Mental Health and Renal Function in Patients With Chronic Kidney Disease: A Randomized Controlled Trial

Hsiao-Ling HUANG[1] • Ya-Hui HSU[2] • Chung-Wei YANG[3] • Min-Fang HSU[4] • Yu-Chu CHUNG[5]*

ABSTRACT

Background: Chronic kidney disease (CKD) refers to permanent damage to the kidneys that occurs gradually over time. Further progression may be preventable depending on its stage.

Purpose: This study was developed to evaluate the effect of a health literacy education program (HLEP) on mental health and renal functioning in patients with CKD.

Methods: A single-blind, randomized controlled trial study was conducted. Data were collected from March 25 to December 18, 2021. Participants were randomly assigned to either the experimental group (n = 42), which received multidisciplinary care and HLEP, or the control group (n = 42), which received multidisciplinary care only. Data were collected at baseline (T1), Month 3 (T2), and Month 6 (T3), and the data included patient characteristics, estimated glomerular filtration rate, and responses to the Mandarin Multidimensional Health Literacy Questionnaire and Beck Depression Inventory.

Results: After 6 months of the HLEP intervention, the results of generalized estimating equations analysis showed that, compared with the control group, the experimental group had significantly higher health literacy at Month 3 (β = −3.37, 95% CI [−5.68, −1.06]), significantly improved depression at Month 3 (β = −2.24, 95% CI [−4.11, −0.37]) and Month 6 (β = −4.36, 95% CI [−6.60, −2.12]), and a significantly higher estimated glomerular filtration rate at Month 6 (β = 5.87, 95% CI [1.35, 10.38]).

Conclusions/Implications for Practice: The findings of this study may provide a reference for healthcare providers to educate patients with Stage 3–4 CKD using the HLEP. Positive effects on health literacy, depression, and renal function in patients with Stage 3–4 CKD were observed in the short term. Findings from this study may facilitate the implementation of multidisciplinary and nurse-led strategies in primary care to reinforce patients' health literacy, self-care ability, and adjustment to CKD as well as delay disease progression.

KEY WORDS:
chronic kidney disease, depression, health literacy, renal function, self-management health education.

Introduction

Chronic kidney disease (CKD) is a global public health issue that affects many people around the world. Of the estimated 14.4% of adults in the United States who have CKD, up to 90% are unaware of their disease (U.S. Renal Data System, 2021). A study on 3,713 patients with Stage 3–4 CKD categorized the sample into four risk levels for renal failure within 5 years: minimum risk (< 2%), low risk (2%–4%), moderate risk (5%–14%), and high risk (≥ 15%). The study concluded that up to 51% of patients in the moderate- and high-risk populations were unaware of their CKD. Patients have a much lower awareness of CKD than of diabetes and hypertension (Chu et al., 2020). However, the onset and progression of a chronic disease tend to be slow, and patients with CKD often have one or more concurrent chronic diseases (Elliott et al., 2020). The progression of chronic disease also affects the mental health of patients. Studies have shown that up to 75.5% of patients with CKD experience depression, which in turn compromises their quality of life (QOL; Kunwar et al., 2020). This emphasizes the importance of aggressive screening and early intervention.

The definition of CKD includes all individuals with an estimated glomerular filtration rate (eGFR) < 60 ml/min per 1.73 m^2 over a 3-month period, an albumin-to-creatinine ratio ≥ 30 mg/g, or other markers of kidney damage regardless of kidney injury status. In those with CKD, the kidneys do not function properly and treatment, for example, kidney

[1]*PhD, Associate Professor, Department of Healthcare Management, Yuanpei University of Medical Technology* • [2]*MSN, RN, Nephrology Case Manager, Department of Internal Medicine, National Taiwan University Hospital Hsin-Chu Branch* • [3]*PhD, MD, Assistant Professor and Attending Physician, Division of Nephrology, Department of Internal Medicine, National Taiwan University Hospital Hsin-Chu Branch* • [4]*PhD, RN, Assistant Professor, Department of Nursing, Yuanpei University of Medical Technology* • [5]*PhD, RN, Professor, Department of Nursing, Yuanpei University of Medical Technology.*

transplantation, is required to avoid kidney failure (National Kidney Foundation, 2022). Therefore, delaying CKD progression is the key focus of medical care. In Taiwan, since the enactment of the National Health Insurance's Healthcare and Health Education for Pre-End Stage Renal Disease Patients Program in 2003, the high-risk population with CKD (patients in Stages 3b–5) has been subject to case management and receives care from a professional medical team responsible for assessing each patient's self-care and self-management abilities. Furthermore, this team follows up on patient health status to help maintain residual renal function. As part of the program, the CKD Clinical Practice Guidance recommends initiation of a low-protein diet (0.6–0.8 g/kg a day) and ketogenic amino acid treatment for patients with Stage 3 CKD to reduce kidney damage caused by nitrogenous wastes and delay dialysis or death (Xu, 2015).

Health literacy is a relatively new field of research in the area of medicine and health. Different age groups require different health literacy programs to empower them to implement beneficial health behaviors (Quaglio et al., 2017). A systematic review by Sørensen et al. (2012) defined health literacy as the "knowledge, motivation, and competencies of accessing, understanding, appraising, and using health-related information within the healthcare, disease prevention and health promotion setting in daily life to make judgment and decisions in order to maintain or improve the overall QOL." Therefore, health literacy influences self-care efficacy and disease prognosis. Lack of health literacy has been reported in 17.7% of patients with Stage 1–5 CKD (Schrauben et al., 2020), and low or absent health literacy has been reported in 22.5% of patients with Stage 3–4 CKD (Hanpaiboon & Pratoomsoot, 2019). Wei et al. (2017) has evaluated the validity of the Mandarin Multidimensional Health Literacy Questionnaire (MMHLQ) on a sample of 2,394 adults in Taiwan, finding the highest score in the domain of "understanding health information," followed by "accessing health information," "communication and interaction," "applying health information," and "appraising health information." Factors that affect health literacy include age, gender, educational level, marital status, spouse cohabitation status, family income, CKD stage, duration of CKD, and number of comorbidities (Y.-C. Chen et al., 2018; Hanpaiboon & Pratoomsoot, 2019; Wong et al., 2018). Patients with higher health literacy show better self-care behaviors (Schrauben et al., 2020).

Patients with CKD tend to experience depression because they are forced to attend numerous hospital visits and face complex treatment plans, drug side effects, dietary restrictions, and uncontrollable clinical symptoms as their disease progresses (S. F. V. Wu et al., 2018). The prevalence of depression is related to CKD stage. A meta-analysis of 22 studies that investigated the correlation between depression and death in patients with CKD reported an average depression prevalence of 27.4% in predialysis patients with CKD (Palmer et al., 2013). Patients with CKD experiencing depression exhibit poor compliance to drug treatment and poor QOL, resulting in increased utilization of medical resources and higher rates of morbidity and mortality (Palmer et al., 2013).

A study conducted by S. F. V. Wu et al. (2018) applied an innovative health education program promoting self-management in a sample of patients with Stage 3b–5 CKD. The program delivered one 100-minute session per week for 4 weeks, and the participants were followed up for 3 months. Outcomes included significantly improved blood urea, nitrogen, and creatinine; reduced depression; and higher self-efficiency and self-management. However, the intervention had no effect on eGFR. Wang et al. (2018) conducted a cross-sectional study to compare the effect of participation in a comprehensive healthcare program on self-care behaviors and kidney function in patients with CKD. The results revealed a slower rate of deterioration in kidney function and better self-management behaviors in patients participating in the healthcare program. Machida et al. (2019) studied the effects of a 1-week inpatient education program on kidney function in patients with Stage 3–5 CKD. The patients were followed from 6 months before hospitalization to 24 months after discharge. Implementation of the program delayed kidney function deterioration during the 2-year observation period, especially in patients with low proteinuria (urinary protein < 0.5 g/gCr). Thus, the authors recommended the program be initiated in patients with low proteinuria. A randomized clinical trial conducted by Lin et al. (2021) investigated patients with Stage 1–3a CKD, with the study group receiving routine care and health coaching for 6 weeks in addition to 12 months of postintervention follow-up. The findings indicate health coaching improves QOL, self-management, patient activation, and self-efficiency.

On the basis of the evidence in the literature, health education programs are beneficial to patients with CKD. Health literacy can influence self-care behaviors and renal function in these patients. However, there is a lack of rigorous research on the impact of health literacy on the psychology of patients with CKD. Therefore, the aim of this study was to investigate patients with Stage 3–4 CKD (the largest group in Taiwan's Pre-End Stage Renal Disease Patients Program), develop a health literacy education program (HLEP), and evaluate the effect of this program on participants' mental health and renal function. This study addressed the following research hypotheses:

1. Patients with CKD who participate in the HLEP will have increased health literacy compared their nonparticipant peers.
2. Patients with CKD who participate in the HLEP will have improved depression compared their nonparticipant peers.
3. Patients with CKD who participate in the HLEP will have improved renal function compared their nonparticipant peers.

Methods

Design

The study design was a single-center, two-group, single-blinded, randomized controlled trial with a repeated-measures design.

The participants were randomized into either the HLEP with multidisciplinary care (experimental) group (EG) or the multidisciplinary care (control) group (CG). Block randomization with 1:1 allocation was conducted using a computer-generated sequence and was performed by one of the authors not involved in screening, patient recruitment, clinical care, or data collection using a random number generator. Sequentially numbered, opaque, sealed envelopes were used to conceal the sequence until the interventions were assigned at an outpatient nephrology clinic. Patients were followed for 6 months. Data were collected before health education (T1) and at 3 months (T2) and 6 months (T3) after completion of the HLEP (Figure 1).

Setting

The participants in this study were conveniently sampled from the nephrology outpatient clinic of a 988-bed regional educational hospital in northern Taiwan.

Participants

Patients who met the selection criteria were recruited. The inclusion criteria were patients aged ≥ 20 years who were able to communicate in Mandarin or Taiwanese, were diagnosed by a nephrologist with Stage 3 or 4 CKD, and had received less than 1 year of comprehensive care. The exclusion criteria included having a cognitive disorder or mental illness (severe depression, schizophrenia), being on routine hemodialysis, or current hospitalization.

Sample Size

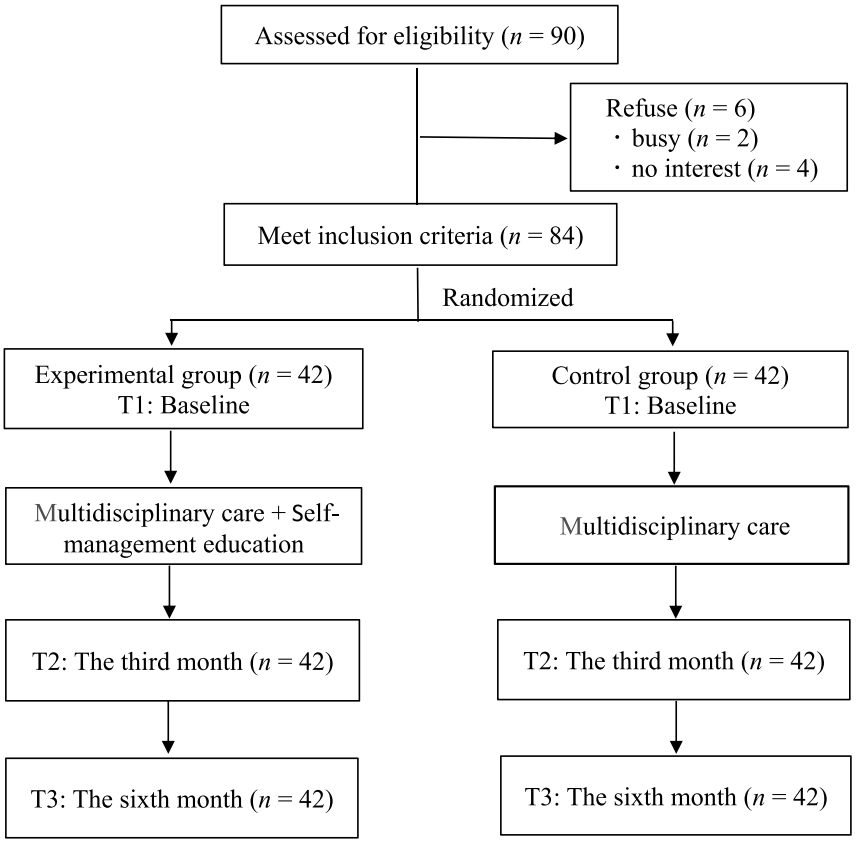
Following Wang et al. (2018), minimal sample size was calculated using G*Power V3.1 statistical software with eGFR as the primary efficacy variable (EG: eGFR = 0.072 ± 8.212, n = 118; CG: eGFR = –2.978 ± 8.680, n = 117). The effect size was estimated as 0.36, α was set at .05, and power was set at 0.95. The participants were divided into two groups with three measurements each. The minimum sample size was calculated as 70. Assuming a follow-up loss of 20%, the final sample size was set as 84 (42 per group).

Experimental Intervention and the Control Group

The main components of the HLEP are shown in Table 1. The HLEP included a self-management health education

Figure 1

Research Design Flowchart



Note. T1: demographic characteristics, Mandarin Multidimensional Health Literacy Questionnaire (MMHLQ), Beck Depression Inventory-II (BDI-II), and estimated glomerular filtration rate (eGFR) were collected. T2 and T3: MMHLQ, BDI-II, and eGFR were collected.

manual and a dietary health education video designed for patients with Stage 3–4 CKD. The health education manual was developed by the researchers based on the Health Literacy Concept and Material Preparation Guide (National Health Insurance Administration, Ministry of Health and Welfare, Taiwan, ROC, 2020) and frequently asked questions from patients and family members. The preliminary review focused on the content and format of the draft, with subsequent revisions made based on comments (Devellis, 2016). Five experts were invited to assess the content validity, with the content validity index assessed in terms of "appropriateness," "accuracy," and "readability" as .98, .88, and .95, respectively, on a 5-point Likert scale, with an overall content validity index of .93. The content of the health education manual was edited based on the experts' comments to create the final version. In the interests of portability, the size of the health education manual was designed as 145 mm in length and 210 mm in width with 15 pages. As most of the participants in this study were older adults, the dietary principles were presented in video format. To maximize learning outcomes, an attending physician from the department of nephrology and one of the researchers personally introduced the dietary principles for patients with Stage 3–4 CKD based on the health education manual. Media professionals were hired to produce the video and sound recordings using PhotoImpact and Adobe Audition for conversion to MP3. The video was designed with a minimum of text and used simple words, pictures, cartoon figures, large fonts, and interactive images.

In the EG, members received one-on-one health education from a study team member with 6 years of experience in kidney disease nursing. The HLEP was delivered using a health education manual in the nephrology outpatient health education classroom. After each session, the participants and their families watched a health education video and were encouraged to ask questions until they fully understood the concepts. In addition, EG participants were taught how to access the videos via their smartphones on YouTube or by scanning a QR code on the cover of the health education manual.

CG participants received routine one-on-one health education from a case manager at the participating hospital who explained the blood analysis results and precautions and distributed an A4-sized health education leaflet.

Data Collection

Data for this study were collected from March 25 to December 18, 2021. The data were collected at three time points: before HLEP implementation (T1) and at 3 months (T2) and 6 months (T3) after HLEP. For participants who were illiterate or had difficulty reading and thus not able to complete the questionnaire independently, a designated staff member explained the questionnaire and assisted them to complete it based on their answers. EG participants received

Table 1

Main Contents of the Health Literacy Education Program

Time	Main Issue	Main Content
Day 1	Chronic kidney disease self-management health education manual (one-on-one health education: 20 minutes)	▪ Introduction of kidney function ▪ Chronic kidney disease and stage ▪ Importance of regular medication ▪ Importance of regular exercise ▪ Blood pressure control ▪ Dietary principles for Stage 3–4 CKD
	Stage 3–4 CKD dietary health education video (10 minutes)	▪ Low protein, enough calories ▪ Dietary principles of potassium restriction ▪ Low-sodium diet ▪ The strategies of blood sugar control and glycated hemoglobin (HbA1c) ▪ Prevent hyperlipidemia ▪ Dietary limitation ▪ Examples of three-meal recipes and substitutions, types of low-protein starches
Week 5, 9, 13, 17, 21	Telephone consultation (5–10 minutes)	▪ Remind patients to watch health education video per month ▪ Discuss video content and dietary principles ▪ Encourage patients to ask questions
Week 12, 24	Routine outpatient follow-up	▪ Discuss health education content and blood test value ▪ Clarify patient concerns

Note. CKD = chronic kidney disease.

multidisciplinary care, participated in the HLEP, and conducted monthly phone discussions with the researcher about the program's content on the first Monday of each month. CG participants received multidisciplinary care, and their data were collected at the same time points as EG participants.

Ethical Considerations

This study was approved by the research ethics committee of National Taiwan University Hospital Hsinchu Branch (Approval No. 91T-27-0026) before initiation. The investigator explained the study purpose and procedure to the participants before they signed informed consent. The participants were informed they were free to withdraw at any time during the study and that their withdrawal would not affect their treatment or cause any negative impact. The study data were coded and analyzed anonymously.

Instruments

Demographic and disease characteristics

The patient characteristics considered in this study included age, gender, educational level, marital status, monthly income, CKD stage, chronic disease history, and duration of treatment in the nephrology department.

Mandarin Multidimensional Health Literacy Questionnaire

The MMHLQ developed specifically for adults in Taiwan by Wei et al. (2017) was used in this study. The 20 items of the MMHLQ assess health information, health information comprehension, health information appraisal, health information application, communication, and interaction using a 4-point scale: 1 = *very difficult*, 2 = *difficult*, 3 = *easy*, and 4 = *very easy*. Total and subscale scores are converted to a 0–50 range using the equation (Mean – 1) × (50/3), with 0–25 indicating *insufficient*, 25–33 indicating *limited*, 33–42 indicating *sufficient*, and 42–50 indicating *excellent* level of health literacy. Higher scores on the MMHLQ indicate better health literacy. The internal consistency reliability analysis revealed good internal consistency (Wei et al., 2017), with a Cronbach's alpha of .92 in this study.

Beck Depression Inventory

The 21-item Beck Depression Inventory-II (BDI-II) Chinese version was used in this study to measure depression. Items are scored on a 4-point Likert scale, with 0 = *no*, 1 = *mild*, 2 = *moderate*, and 3 = *severe* symptoms. The participant chooses the statement in each item that best describes how they felt over the past 2 weeks (including the day of the examination). The total score ranges between 0 and 63, with 0–13 indicating *normal emotion*, 14–19 indicating *mild depression*, 20–28 indicating *moderate depression*, and 29–63 indicating *severe depression*. The BDI-II is aligned with the diagnostic principles of depression in the *Diagnostic and Statistical Manual Disorders, Fourth Edition* and thus may be used to determine depression status. The BDI-II has excellent validity (H. Y. Chen, 2000), and the internal consistency Cronbach's α value in this study was .82.

Renal function

In this study, eGFR was used to monitor kidney function. Data were collected from medical records, and eGFR was calculated using the Modification of Diet in Renal Disease simplified equation developed by the Modification of Diet in Renal Disease Study Group (Levey et al., 2007). The participating hospital scheduled visits for the patients every 3 months, and one blood sample was collected based on the guidelines of the comprehensive care program for kidney disease. Blood samples were collected from the patient during the week before the scheduled visit date.

Data Analysis

IBM SPSS Statistics for Windows 20.0 (IBM Inc., Armonk, NY, USA) was used for data archiving and statistical analysis, with results presented as frequency, percentage, mean, and standard deviation. The demographic and disease characteristics were compared between the groups using the χ^2 test, independent samples t test, and paired samples t test. The outcome variables, including health literacy, depression, and eGFR, were compared between the two groups using a generalized estimating equation with repeated measures.

Results

Eighty-four participants completed the study (0% attrition), with 42 each in the EG and CG. Age ranged from 30 to 87 years and averaged 65.39 (*SD* = 11.39) years. No significant differences between the two groups were found in terms of demographic and disease characteristics (Table 2).

As shown in Table 3, no significant differences between the two groups were found in terms of health literacy, depression, or eGFR before the HLEP intervention. The average MMHLQ score of the participants before health education was categorized as "limited." The highest scores for both groups were reported in the domain of "understanding health information," followed by "communication and interaction," "applying health information," "appraising health information," and "accessing health information." In terms of depression, only 16 (19.1%) participants reported depression before the intervention, including 15 (17.9%) with mild and one (1.2%) with severe depression. There were 31, 10, 0, and 1 participant in the EG and 37, 5, 0, and 0 participants in the CG with normal, mild, moderate, and severe depression, respectively, with no significant between-group difference in depression noted (χ^2 = 10.129, p = .928). The MMHLQ, depression, and eGFR scores over time for the two groups are presented in Figure 2, and a summary of the GEE results for MMHLQ, depression, and eGFR is shown in Table 4. A model with an exchangeable correlation matrix and model-based estimates of variance was used.

Table 2

Homogeneity Test of Demographic and Clinical Characteristics (N = 84)

Variable	Overall (*N* = 84)		EG (*n* = 42)		CG (*n* = 42)		χ^2	*p*
	n	%	*n*	%	*n*	%		
Gender							0.66	.588
Female	17	20.2	10	23.8	7	16.7		
Male	67	79.8	32	76.2	35	83.3		
Age, years							2.30	.512
≤ 64	35	41.7	18	42.9	17	40.5		
65–74	33	39.3	16	38.1	17	40.5		
≥ 75	16	19.0	8	19.0	8	19.0		
Education							2.69	.260 [a]
High school and below	55	65.5	24	57.1	31	73.8		
Junior college	9	10.7	6	14.3	3	7.1		
University and above	20	23.8	12	28.6	8	19.1		
Marital status							1.40	.433 [a]
Not married	7	8.3	2	4.8	5	11.9		
Married	77	91.7	40	95.2	37	88.1		
Monthly income (NT$)							0.75	.119 [a]
No	43	51.2	17	40.5	26	61.9		
≤ 20,000	12	14.3	5	11.9	7	16.6		
20,001–39,999	12	14.3	7	16.7	5	11.9		
40,000–49,999	4	4.8	4	9.5	0	0.0		
50,000–59,999	8	9.5	6	14.3	2	4.8		
≥ 60,000	5	5.9	3	7.1	2	4.8		
CKD stage							1.93	.381
3a	18	21.4	8	19.0	10	23.8		
3b	38	45.2	17	40.4	21	50.0		
4	28	33.4	17	40.5	11	26.2		
Chronic disease history							1.64	.442 [a]
1 category	21	25.0	13	31.0	8	19.1		
2 categories	36	42.9	17	40.4	19	45.2		
≥ 3 categories	27	32.1	12	28.6	15	35.7		
Hypertension							0.66	.415 [a]
No	17	20.2	10	23.8	7	16.7		
Yes	67	79.8	32	76.2	35	83.3		
Diabetes							1.20	.274
No	39	46.4	22	52.4	17	40.5		
Yes	45	53.6	20	47.6	25	59.5		
Gout							0.00	1.000
No	54	64.3	27	64.3	27	64.3		
Yes	30	35.7	15	35.7	15	35.7		
Albumin (g/dl)							17.71	.314
Normal (3.5–5.7)	81	96.4	41	97.6	40	95.2		
Abnormal (< 3.5, > 5.7)	3	3.6	1	2.4	2	4.8		
Proteinuria (g/dl)							80.00	.447
Normal (≤ 150)	23	27.6	9	21.6	14	33.6		
Abnormal (> 150)	61	72.4	33	78.4	28	66.4		

Note. EG = experimental group; CG = control group; CKD = chronic kidney disease.

[a] Fisher's exact test.

After adjusting for age and gender, relationships between health literacy and, respectively, time, group, and Time × Group interaction were explored. The model showed the time effect as more significant for T2 and T3 compared with T1. Trend differences (interactions between time and group) revealed significant differences in health literacy at T2,

Table 3

Homogeneity Test of MMHLQ, Depression, and eGFR (N = 84)

Variable	EG (*n* = 42)		CG (*n* = 42)		*t*	*p*
	Mean	*SD*	Mean	*SD*		
MMHLQ/overall	30.06	5.78	28.06	7.27	−1.40	.166
Accessing health information	25.60	11.55	23.61	12.47	−0.76	.452
Understanding health information	35.12	6.96	34.92	8.73	−0.12	.909
Appraising health information	27.27	6.96	24.70	8.60	−1.51	.135
Applying health information	29.66	7.71	26.69	10.62	−1.47	.145
Communication and interaction	32.64	6.03	30.36	6.39	−1.68	.096
Depression	9.29	6.99	7.12	4.72	−1.67	.100
eGFR (ml/min per 1.73 m^2)	34.76	12.69	37.16	10.27	0.88	.384

Note. MMHLQ = Mandarin Multidimensional Health Literacy Questionnaire; eGFR = estimated glomerular filtration rate; EG = experimental group; CG = control group.

indicating time-dependent growth effects. Trend differences revealed significant improvements in depression scores for the EG at T2 and T3, indicating time-dependent growth effects. In terms of eGFR, the time effect was more significant at T3 compared with T1. Trend differences revealed a significant increase in eGFR for the EG at T3, indicating time-dependent growth effects. The overall mean eGFR at pretest (T1) was 35.96 (*SD* = 11.53) ml/min per 1.73 m^2 and, at T3, had increased in the EG from 34.96 ml/min per 1.73 m^2 to 38.83 ml/min per 1.73 m^2 and decreased in the CG from 37.16 ml/min per 1.73 m^2 to 35.35 ml/min per 1.73 m^2.

Discussion

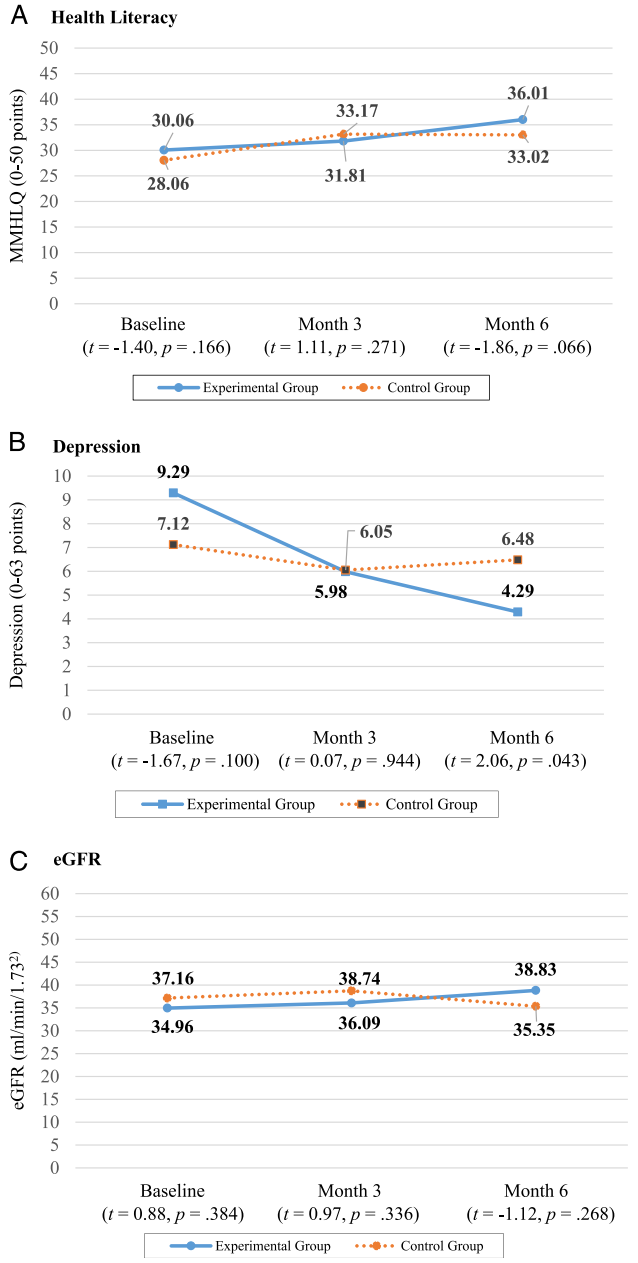
The participant characteristics align with the general population with Stage 3–4 CKD in Taiwan, with a predominantly male (79.8%) and Stage 3 CKD (66.6%) study group, averaging 65.39 years old (*SD* = 11.39). According to statistics from the 2016 to 2018 kidney disease in Taiwan annual report, CKD is most prevalent in older adult men, with Stage 3 being the most commonly observed stage (National Health Research Institutes & Taiwan Social of Nephrology, 2020). The top three concomitant chronic diseases reported by the participants were hypertension (79.8%), diabetes (53.0%), and gout (35.7%). The largest percentage (40.5%) had two concomitant diseases, 32.1% had three, and 25% had one. These findings are consistent with previous statistical data on patients with CKD, who frequently report one to two concomitant chronic diseases (Elliott et al., 2020). According to the 2020 kidney disease in Taiwan annual report, the top three concomitant diseases reported in the previous year by new dialysis patients were hypertension, heart disease, and diabetes (National Health Research Institutes & Taiwan Social of Nephrology, 2020). These data were based on the kidney biopsy results presented in the kidney disease in Taiwan annual report, in which nephrotic syndrome and proteinuria of unknown cause accounted for 46.4% of the cases, and 20.2% of acute renal injury was caused by the improper use of nonsteroidal anti-inflammatory drugs. Nonsteroidal anti-inflammatory drugs are the most used drugs by patients with gout (National Health Research Institutes & Taiwan Social of Nephrology, 2020).

The health literacy of the participants before the intervention was "limited," which is similar to the result reported by the National Health Research Forum (2020). Up to 50% of older patients lack sufficient health information for health decision making because of an insufficient level of health literacy. The MMHLQ used in this study returned the highest score for both groups in the "understanding health information" domain, followed by "appraising health information" and "accessing health information." Only the "understanding health information" domain achieved a "sufficient" level, whereas the other domains were "limited." This finding differs from the results of several previous studies (Wei et al., 2017; C. L. Wu et al., 2020). Wei et al. reported the highest mean MMHLQ score in the "understanding health information" domain, followed by the "accessing health information," "communication and interaction," "applying health information," and "appraising health information" domains. C. L. Wu et al., investigating the health literacy of 458 patients from eight patient support groups, reported the highest mean MMHLQ score in the "understanding health information" domain, followed by the "communication and interaction," "accessing health information," "applying health information," "understanding health information," and "appraising health information" domains. The difference in findings may be attributed to differences in study purposes and samples. Nevertheless, health education materials and approaches for patients with CKD should emphasize these four domains of MMHLQ, with particular focus given to the health information domain.

Health literacy improved significantly in the EG between T1 and T2 and remained relatively unchanged between T2 and T3. One possible reason for this outcome is that 65.5% of the participants had an educational level below

Figure 2

Group Comparisons of Outcomes at Baseline, Month 3, and Month 6



Note. MMHLQ = Mandarin Multidimensional Health Literacy Questionnaire; eGFR = estimated glomerular filtration rate.

the college level and were predominantly older individuals who likely required more time to learn. The level of health literacy before health education was "limited" for both groups and had improved to "sufficient" for the EG at T2 and T3 and to "sufficient" for the CG only at T3. The EG showed a more significant growth trend in health literacy compared with the CG, indicating that an even longer follow-up period is necessary to determine the long-term effects of the intervention.

Before the intervention, only 19.1% of the participants reported experiencing depression, with 17.9% having mild depression and 1.2% having severe depression. This differs from Loosman et al. (2015), which found a 34% prevalence of depression among patients with Stage 3b–4 CKD and a mean eGFR of 20.4 (*SD* = 6.3) ml/min per 1.73 m^2 at baseline. Their findings suggest that the proportion of patients with depression rises as renal function declines. The prevalence of depression among the participants in this study was relatively low, which may be explained by most participants (65.5%) being in Stage 3 CKD and the overall mean eGFR at baseline being 35.96 ± 11.53 ml/min per 1.73 m^2, indicating better renal function. At T2 and T3, depression had improved significantly in the EG, which is similar to S. F. V. Wu et al. (2018), who also employed an innovative self-management program on a group of patients with Stage 3b–5 CKD. In their study, depression decreased significantly at 3 months after the intervention, which included individual health education every 3 months during the study period and a visit and phone interview on the first week of each month. Regular contact and establishing excellent rapport helped make subjects feel concern for their physical and mental issues, which further improved their depression.

The participants in this study included 65.5% in Stage 3 and 34.5% in Stage 4 CKD. According to the CKD prevalence in adults in a cohort study conducted from 2015 to 2018, 5.8% patients had Stage 3, 0.4% had Stage 4, and 0.1% had Stage 5 CKD, with Stage 3 comprising the highest proportion (U.S. Renal Data System, 2021). In our study, after the intervention, eGFR increased consistently with time in the EG and decreased significantly with time in the CG until T3, at which time the eGFR was significantly higher in the EG than the CG. This result supports the recommendations of S. F. V. Wu et al. (2018). In their study, the EG was enrolled in an innovative self-care program, and no significant improvement was observed in eGFR after 3 months of follow-up. The authors suggested that the follow-up period should be extended to 6–12 months to better detect the effect. Wang et al. (2018) reported similar results in their study investigating patients with Stage 1–5 CKD. In patients receiving more than 1 year of comprehensive CKD care, the healthcare program group attained a 2.83-fold higher likelihood of experiencing a slower deterioration of kidney function than the non-healthcare program group. eGFR increased by 3.87 ml/min per 1.73 m^2 on average in the EG and decreased by 1.81 ml/min per 1.73 m^2 on average in the CG at 6 months after the intervention. On the basis of their findings, HLEP was deployed in this study to enhance kidney function.

Limitations of the Study

This study was affected by several limitations. First, recruitment in this study was limited to patients with Stage 3–4 CKD at a single district teaching hospital in northern Taiwan. Although their characteristics were similar to the

Table 4

Results of Generalized Estimating Equations for MMHLQ, Depression, and eGFR Score (N = 84)

Parameter	MMHLQ			Depression			eGFR		
	β	95% CI	*p*	β	95% CI	*p*	β	95% CI	*p*
Intercept	42.48	[32.92, 52.05]	< .001	4.01	[−1.70, 9.73]	.170	74.29	[65.90, 82.68]	< .001
Experimental group [a]	2.14	[−0.28, 4.56]	.082	1.86	[−0.60, 4.32]	.138	0.28	[−1.62, 2.18]	.770
Month 3 (T2) [b]	5.95	[3.85, 8.06]	< .001	−1.07	[−2.23, 0.09]	.070	1.58	[−0.10, 3.26]	.060
Month 6 (T3) [b]	5.79	[4.04, 7.55]	< .001	−0.64	[−2.14, 0.85]	.400	−1.80	[−3.40, −0.20]	.030
Experimental Group × T2 [c]	−3.37	[−5.68, −1.06]	.004	−2.24	[−4.11, −0.37]	.019	−0.25	[−3.15, 2.65]	.870
Experimental Group × T3 [c]	0.99	[−1.42, 3.40]	.419	−4.36	[−6.60, −2.12]	< .001	5.87	[1.35, 10.38]	.011

Note. Adjusted: age and gender. MMHLQ = Mandarin Multidimensional Health Literacy Questionnaire; eGFR = estimated glomerular filtration rate.
[a] Reference group: control group. [b] Reference group: baseline (T1). [c] Reference group: Control Group × Baseline.

general population of patients with Stage 3–4 CKD in Taiwan, the results may not be generalizable to patients with CKD nationwide. Future studies should include larger sample sizes from multiple hospitals with different stages of CKD. Second, this study followed the patients for a postintervention period of 6 months only, which may not capture the long-term effects on health literacy of the HLEP program. Therefore, the study period should be extended in future studies to evaluate long-term efficacy. Third, the participants in this study were patients who had participated in a multidisciplinary care program for less than a year. Future studies should include patients with no prior multidisciplinary care program experience to determine the effectiveness of the self-management health education program on this patient group. Finally, the study was conducted in a single clinic, and the intervention and data collection were performed by the same researcher, which may have introduced a Hawthorne effect. To minimize the potential for this bias, future studies should recruit from multiple clinics and use different researchers for intervention and data collection tasks.

Conclusions and Implications for Practice

The results of this study indicate the developed HLEP program significantly and positively affects the health literacy of patients with CKD within 3 months of program completion. Moreover, at 6 months posttest, the severity of depression decreased and kidney function improved significantly in the EG. Furthermore, the mean eGFR increased in the EG by 3.87 ml/min per 1.73 m^2 at 6 months posttest, whereas eGFR decreased by 1.81 ml/min per 1.73 m^2 in the CG. These findings suggest the HLEP may be an effective tool for improving health literacy and clinical outcomes in patients with CKD. This study may serve as a reference for nephrology case managers responsible for educating patients with Stage 3–4 CKD using the HLEP. In practice, the HLEP developed in this study may be accessed by patients at home on their computer/cellphone or via the provided QR code. The provided health education manual is small and easy to carry and may be used as a reference when dining out. Health education manuals are a more convenient format than traditional health education leaflets. A good health education tool can enhance case manager performance, increase patient confidence in controlling their kidney disease, and delay disease progression. Therefore, this tool is worth considering in practice. Overall, the results of this study suggest providing tailored health education to patients with Stage 3–4 CKD using the HLEP can improve health literacy and clinical outcomes and benefit both patients and healthcare providers.

Author Contributions

Study conception and design: YCC, HLH
Data collection: YHH
Data analysis and interpretation: YCC, YHH, CWY, MFH
Drafting of the article: YCC, YHH
Critical revision of the article: YCC, HLH

Received: October 1, 2022; Accepted: May 19, 2023
*Address correspondence to: Yu-Chu CHUNG, PhD, RN, Department of Nursing, Yuanpei University of Medical Technology, No. 306, Yuanpei Street, Hsinchu City 30015, Taiwan, ROC. Tel: +886-3-6102309; E-mail: yuchu@mail.ypu.edu.tw
The authors declare no conflicts of interest.

Cite this article as:
Huang, H.-L., Hsu, Y.-H., Yang, C.-W., Hsu, M.-F., & Chung, Y.-C. (2024). Effects of a health literacy education program on mental health and renal function in patients with chronic kidney disease: A randomized controlled trial. *The Journal of Nursing Research, 32*(1), Article e310. https://doi.org/10.1097/jnr.0000000000000595

References

Chen, H. Y. (2000). *Manual for the Beck Depression Inventory-II*. Chinese Behavioral Science. (Original work published in Chinese)

Chen, Y.-C., Chang, L.-C., Liu, C.-Y., Ho, Y.-F., Weng, S.-C., & Tsai, T.-I. (2018). The roles of social support and health literacy in self-management among patients with chronic kidney disease. *Journal of Nursing Scholarship, 50*(3), 265–275. https://doi.org/10.1111/jnu.12377

Chu, C. D., McCulloch, C. E., Banerjee, T., Pavkov, M. E., Burrows, N. R., Gillespie, B. W., Saran, R., Shlipak, M. G., Powe, N. R., Tuot, D. S., & Centers for Disease Control and Prevention Chronic Kidney Disease Surveillance Team. (2020). CKD awareness among US adults by future risk of kidney failure. *American Journal of Kidney Diseases*, *76*(2), 174–183. https://doi.org/10.1053/j.ajkd.2020.01.007

Devellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Sage Publications.

Elliott, M. J., Love, S., Donald, M., Manns, B., Donald, T., Premji, Z., Hemmelgarn, B. R., Grinman, M., Lang, E., & Ronksley, P. E. (2020). Outpatient interventions for managing acute complications of chronic diseases: A scoping review and implications for patients with CKD. *American Journal of Kidney Diseases*, *76*(6), 794–805. https://doi.org/10.1053/j.ajkd.2020.04.006

Hanpaiboon, K., & Pratoomsoot, C. (2019). Factors influencing patient health behaviors for delaying the progress in stage 3–4 chronic kidney disease patients at Khlongkhlung hospital, Khampangphet province. *Thai Pharmaceutical and Health Science Journal*, *14*(2), 53–61.

Kunwar, D., Kunwar, R., Shrestha, B., Amatya, R., & Risal, A. (2020). Depression and quality of life among the chronic kidney disease patients. *Journal of Nepal Health Research Council*, *18*(3), 459–465. https://doi.org/10.33314/jnhrc.v18i3.2556

Levey, A. S., Coresh, J., Greene, T., Marsh, J., Stevens, L. A., Kusek, J. W., & Van Lente, F.; Chronic Kidney Disease Epidemiology Collaboration. (2007). Expressing the modification of diet in renal disease study equation for estimating glomerular filtration rate with standardized serum creatinine values. *Clinical Chemistry*, *53*(4), 766–772. https://doi.org/10.1373/clinchem.2006.077180

Lin, M. Y., Cheng, S. F., Hou, W. H., Lin, P. C., Chen, C. M., & Tsai, P. S. (2021). Mechanisms and effects of health coaching in patients with early-stage chronic kidney disease: A randomized controlled trial. *Journal of Nursing Scholarship*, *53*(2), 154–160. https://doi.org/10.1111/jnu.12623

Loosman, W. L., Rottier, M. A., Honig, A., & Siegert, C. E. (2015). Association of depressive and anxiety symptoms with adverse events in Dutch chronic kidney disease patients: A prospective cohort study. *BMC Nephrology*, *16*(1), Article No. 155. https://doi.org/10.1186/s12882-015-0149-7

Machida, S., Shibagaki, Y., & Sakurada, T. (2019). An inpatient educational program for chronic kidney disease. *Clinical and Experimental Nephrology*, *23*(4), 493–500. https://doi.org/10.1007s10157-018-1660-5

National Health Insurance Administration, Ministry of Health and Welfare, Taiwan, ROC. (2020). *Pre-ESRD patient care and health education program*. https://www.nhi.gov.tw/Content_List.aspx?n=74FB9F36D1234D73&topn=5FE8C9FEAE863B46 (Original work published in Chinese)

National Health Research Forum. (2020). *The determinates of health literacy and related health outcomes among elders in Taiwan*. https://forum.nhri.edu.tw/wp-content/uploads (Original work published in Chinese)

National Health Research Institutes & Taiwan Social of Nephrology. (2020). *2020 kidney disease in Taiwan annual report*. Authors.

National Kidney Foundation. (2022). *Chronic kidney disease (CKD) symptoms and causes*. https://www.kidney.org/atoz/content/about-chronic-kidney-disease#chronic-kidney-disease-covid-19

Palmer, S. C., Vecchio, M., Craig, J. C., Tonelli, M., Johnson, D. W., Nicolucci, A., Pellegrini, F., Saglimbene, V., Logroscino, G., Hedayati, S. S., & Strippoli, G. F. (2013). Association between depression and death in people with CKD: A meta-analysis of cohort studies. *American Journal of Kidney Disease*, *62*(3), 493–505. https://doi.org/10.1053/j.ajkd.2013.02.369

Quaglio, G., Sørensen, K., Rübig, P., Bertinato, L., Brand, H., Karapiperis, T., Dinca, I., Peetso, T., Kadenbach, K., & Dario, C. (2017). Accelerating the health literacy agenda in Europe. *Health Promotion International*, *32*(6), 1074–1080. https://doi.org/10.1093/heapro/daw028

Schrauben, S. J., Cavanaugh, K. L., Fagerlin, A., Ikizler, T. A., Ricardo, A. C., Eneanya, N. D., & Nunes, J. W. (2020). The relationship of disease-specific knowledge and health literacy with the uptake of self-care behaviors in CKD. *Kidney International Reports*, *5*(1), 48–57. https://doi.org/10.1016/j.ekir.2019.10.004

Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., Brand, H., & (HLS-EU) Consortium Health Literacy Project European. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*, *12*(1), Article No. 80. https://doi.org/10.1186/1471-2458-12-80

U.S. Renal Data System. (2021). *2021 Annual data report*. https://adr.usrds.org/2021/chronic-kidney-disease

Wang, S.-L., Chen, T.-H., Kung, L.-F., Chiou, C.-L., & Lin, M.-Y. (2018). The effect of multidisciplinary integrated care program on self-care behavior and renal function in patient with chronic kidney disease. *Chang Gung Nursing*, *29*(1), 1–15. https://doi.org/10.6386/CGN.201803_29(1).0001 (Original work published in Chinese)

Wei, M.-H., Wang, Y.-W., Chang, M.-C., & Hsieh, J.-G. (2017). Development of Mandarin Multidimensional Health Literacy Questionnaire (MMHLQ). *Taiwan Journal of Public Health*, *36*(6), 556–570. https://doi.org/10.6288/TJPH201736106061 (Original work published in Chinese)

Wong, K. K., Velasquez, A., Powe, N. R., & Tuot, D. S. (2018). Association between health literacy and self-care behaviors among patients with chronic kidney disease. *BMC Nephrology*, *19*(1), Article No. 196. https://doi.org/10.1186/s12882-018-0988-0

Wu, C.-L., Liou, C.-H., Liu, S.-A., Sheu, W.-H., & Tsai, S.-F. (2020). Mandarin Multidimensional Health Literacy Questionnaire for patient supporting groups: A quality improvement article. *Medicine*, *99*(45), Article e23182. https://doi.org/10.1097/MD.0000000000023182

Wu, S. F. V., Lee, M. C., Hsieh, N. C., Lu, K. C., Tseng, H. L., & Lin, L. J. (2018). Effectiveness of an innovative self-management intervention on the physiology, psychology, and management of patients with pre-end-stage renal disease in Taiwan: A randomized, controlled trial. *Japan Journal of Nursing Science*, *15*(4), 272–284. https://doi.org/10.1111/jjns.12198

Xu, Z.-C. (2015). *Taiwan chronic kidney disease clinical guidelines —2015*. National Health Research Institutes. (Original work published in Chinese)

Gut Microbiota and Probiotics in Health and Disease

This webinar took place on 23rd September 2021, as part of the 5th Global Microbiota Summit

Speakers: Mary Ellen Sanders,[1] Ana Teresa Abreu,[2] Karine Clément[3]

1. Dairy and Food Culture Technologies, Centennial, Colorado, USA
2. Hospital Ángeles del Pedregal, Mexico City, Mexico
3. Inserm/Sorbonne University, Pitié-Salpêtrière Hospital, Paris, France

Disclosure: Sanders has received consultancy and speaker fees from Bayer, Bloom Pharmaceuticals, The Chronic Disease Research Foundation (CDRF), Church & Dwight, PepsiCo, Kerry, Associated British Foods, Mead Johnson, Fairlife, GlaxoSmithKline, Trouw Nutrition, Allergosan OMNi-BiOTiC, Probi, Sanofi, Cargill, Danone North America, Danone Research, Winclove Probiotics, and Yakult. Abreu has received consultancy, speaker, or research fees from Sanofi, AB Biotics, Axon Pharma, Mayoly Spindler, Biocodex, Alfasigma, Tecnoquímicas, MD Pharma, Medix Healthcare, Menarini, Ferrer, Takeda, Carnot (Mexico), Adare Pharmaceuticals, Abbott, Faes Farma, Falk Institute, Instituto de Nutrición y Salud Kellogg's, and Danone Institute. Clément has received consultancy, speaker, or research fees from Danone Research, Ysopia, Confo Therapeutics, and Sanofi Consumer HealthCare.

Acknowledgements: Writing assistance was provided by Nicola Humphry, Nottingham, UK.

Support: The symposium and publication of this article were funded by Sanofi CHC. The views and opinions expressed are exclusively those of the speakers. The content was reviewed by Sanofi CHC for medical accuracy.

Citation: EMJ Gastroenterol. 2021;10[1]:42-50.

Meeting Summary

Mary Ellen Sanders opened the webinar by defining and differentiating the 'biotic' family, including probiotics, prebiotics, synbiotics, and postbiotics. She discussed the need for improved labels on commercial products in the biotics family and emphasised the research gaps in this field Ana Teresa Abreu expanded on a specific probiotic, *Bacillus clausii*, describing the evidence for health benefits associated with this bacterium and the potential mechanisms through which it might achieve these effects. Finally, Karine Clément discussed the role of the gut microbiome in cardiometabolic disease, suggesting that gut microbiota may represent a missing link between the environmental and genetic factors that impact these diseases. Clément described the evidence for a dysbiosis of gut microbiota in metabolic diseases and posited that a personalised approach to gut microbiome therapy might be the best way to leverage this association.

The Science of Probiotics and Related Biotics: How to Understand and Use Them

Mary Ellen Sanders

Mary Ellen Sanders introduced 'biotics' as a family of four microbiome-targeted substances: probiotics, prebiotics, synbiotics, and postbiotics. Each type of biotic has the potential to impact the resident microbes of a host, which have diverse physiological functions, including promotion of fat storage and angiogenesis, immune development, synthesis of vitamins and amino acids, drug metabolism, modification of the nervous system, breakdown of food, resistance to pathogens, protection against epithelial injury, and modulation of bone-mass density.[1]

Many human diseases and disorders are associated with an altered microbiome, including irritable bowel disease, colon cancer, diabetes, obesity, rheumatoid arthritis, and liver disease.[1–3] However, Sanders emphasised that it is not yet clear whether the altered microbiome is a cause or a result of these conditions.[1] This raises the question of whether restoring the microbiota in individuals with these conditions, to match that of healthy individuals, would affect the condition itself.

Biotics are intended to influence colonising microbiota to improve health, but understanding of what constitutes a healthy microbiome is still quite limited. Sanders explained that rather than focusing on the specific microbes present in the microbiome, a healthy microbiome may be better characterised by a high diversity of taxonomic units, high resilience (the ability to recover from perturbations such as antibiotic exposure), and functional redundancy (more than one ecosystem member can perform the same function).[4]

Sanders feels that although studying the microbiome is helpful to understand the mechanisms of biotics, the evidence of health benefits is more important. For example, probiotics have been shown to benefit health for various clinical endpoints, across the human lifespan, and in different organ systems, such as preventing antibiotic-associated or traveller's diarrhoea, treating ulcerative colitis, and reducing the incidence of infection gastrointestinal disease.[2] For most of these benefits, a microbiome-mediated mechanism has not been demonstrated yet.[4]

The International Scientific Association for Probiotics and Prebiotics (ISAPP) has published statements that include clear definition for each of member of the biotic family, based on consensus panels (Table 1). Importantly, these definitions are deliberately broad enough to support innovation; they do not restrict these substances by host (e.g., human, agricultural animals, etc), regulatory category (e.g., food,

Table 1: ISAPP Definitions of biotic substances.

Biotic substance	Definition
Probiotic	Live microorganisms that, when administered in adequate amounts, confer a health benefit on the host[5]
Prebiotic	A substrate that is selectively utilised by host microorganisms, conferring a health benefit on the host[6]
Synbiotic	A mixture comprising live microorganisms and substrate(s) selectively utilised by host microorganisms that confers a health benefit on the host[7]
Postbiotic	Preparation of inanimate microorganisms and/or their components that confers a health benefit on the host[8]

Definitions are concise; for full understanding, see the full statements. All substances must be safe for their intended use.

ISAPP: International Scientific Association for Probiotics and Prebiotics.

drug, or supplements), site of action (e.g., gut, vaginal tract, skin, etc), or mechanism of action.[5-8]

Probiotics

A number of different microbes are used as probiotics, many of which are members of the *Lactobacillaceae* family or are species of *Bifidobacterium*, *Bacillus,* or *Saccharomyces.*[9] The range of probiotic species is rapidly expanding[10] as more is learnt about the microbes that reside in the healthy human body. Sanders emphasised the importance of recognising that probiotics are a heterogenous group; two products which contain the same microbial genus and species but differ by microbial strain may differ in function.

To be defined as a probiotic, a substance must be a properly identified (both sequenced and named). The microbe must be alive when administered, and studies need to have demonstrated a health benefit for a specific target host at the specific dose delivered by the product. In addition, the microbial strain and manufacturing process must be safe for the intended use, and the product must be correctly labelled with the strain and colony forming units (CFU) expected at the end of its shelf life.[5]

Ideally, probiotic product labels should detail health benefits (supported by evidence), suggested serving size, proper storage conditions, and contact details for consumer information.[11] However, a survey of refrigerated probiotic foods in grocery stores in the USA found that only one-half (22 of 45) of products listed the constituent microbial strains. Those that did, could be linked to evidence of health benefits, tended to contain fewer strains, and had a lower CFU per serving compared to other products.[12] A survey of probiotic supplements found similar results: most products could not be linked to evidence; 45% did not list constituent microbial strains; and 45% did not provide CFU at end of shelf life.[13] Sanders emphasised that similar problems exist outside of the USA.

Neither probiotics nor postbiotics are required to target the microbiome directly, whereas prebiotics and synbiotics should do so as part of their mechanism of action.[5-8] Despite these distinctions, Sanders explained that there is a common belief among both scientists and the general public that probiotics have an important impact on the gut microbiome. This belief is not fully substantiated by the available research data; a systematic review of clinical trials showed that probiotics did not have a global impact on the faecal microbial communities in healthy subjects.[14] Sanders suggested that this does not prove that probiotics have no effect; their effects may be limited to minor components of the microbiota, not evident in faecal samples or in healthy subjects, or only evident in the metabolites rather than the microbiome composition. However, she stressed that the evidence to date indicates that the effects of current probiotics on the microbiome are likely to be quite subtle.

In summary, Sanders reiterated that the healthy gut microbiome has not yet been defined by researchers, but that for probiotics, effects on the microbiome are probably less important than health benefits. There is a clear need for improved labels on commercial products in the biotics family so that healthcare practitioners and consumers know what they are buying, and the terms probiotics, prebiotics, synbiotics, and postbiotics should only be used when the scientifically accepted criteria are met. She emphasised the research gaps in this arena, including defining a healthy microbiome, robust trials to confirm health benefits, and identifying the best strains and doses for specific applications. Finally, Sanders emphasised that it will be important to understand the mechanisms that drive the clinical benefits of biotics in order to optimise these substances for future use.

Bacillus clausii: Mechanisms as Spore Probiotics in Gastrointestinal Disorders

Ana Teresa Abreu

Bacillus is one of the most studied bacterial genera[15] and its species can be found in soil, water, food, and in the human gut.[16] These aerobic bacteria can differentiate into a dormant endospore, allowing them to survive in stomach acid and bile salts in the gastrointestinal system.[16,17]

Most *Bacilli* are not pathogenic to humans or animals, and in the case of *B. clausii* (Figure 1),[18] an endosymbiotic relationship, where one organism lives inside the other, between

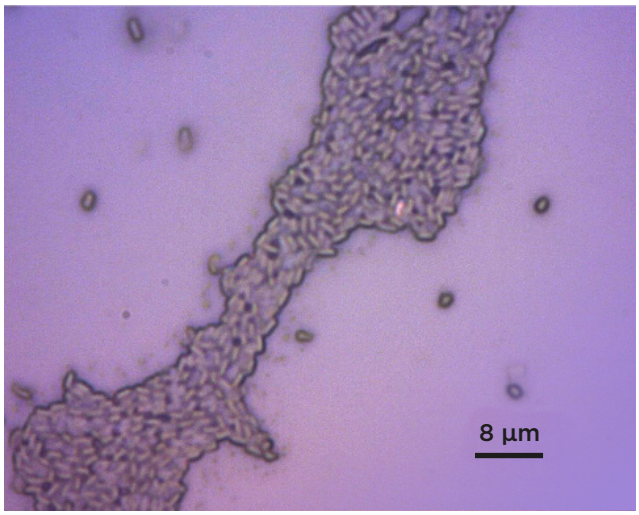


Figure 1: *B. clausii* (combined antibiotic resistant strains: O/C, SIN, N/R, T).
N/R: novobiocin and rifampicin; O/C: chloramphenicol; SIN: streptomycin and neomycin; T: tetracycline.

species and their hosts has been suggested.[17,19,20] Four strains of *B. clausii* are resistant to antibiotics, a property considered advantageous to restoring a healthy gut, and are named for their predominant antibiotic resistance: novobiocin and rifampicin (strain N/R), chloramphenicol (strain O/C), streptomycin and neomycin (strain SIN), and tetracycline (strain T).[19]

There are several properties of *B. clausii* that contribute to its probiotic effects:

- *B. clausii* spores can survive the hostile environment of the gastrointestinal tract and multiply to colonise the intestine.[21-23]
- The pan-genome of *B. clausii* (O/C, SIN, N/R, T) includes genes involved in carbohydrate metabolism,[19] and one strain, SKAL 16, has been shown to excrete butyrate in *in vitro* conditions.[24] Butyrate serves as the major energy source for enterocytes, exerts anti-inflammatory effects, and enhances gut barrier function.[25]
- The antibiotic resistance genes of *B. clausii* are stable and cannot be transferred to other bacteria.[26] Many strains of *B. clausii* are recommended for use along with antibiotics, and Abreu emphasised that it is important for clinicians to match probiotic strains to the prescribed antibiotic therapy.
- Some strains of *B. clausii,* particularly SIN and T, produce the essential vitamin riboflavin (vitamin B2) *in vitro*, suggesting that *B. clausii* has the potential to compensate for host deficits in riboflavin that can occur in clinical contexts such as chemotherapy.[27]
- *Bacillus* species produce a wide range of antimicrobial substances, including lantibiotics (post-translationally modified peptides) which are active against gram-positive bacteria such as *Clostridium difficile.*[20,28] One such lantibiotic, clausin, has been isolated from *B. clausii* and interacts with lipid intermediates in the bacterial envelope biosynthesis pathways,[29] suggesting that it could help to manipulate the constituents of the intestinal microbiota.

Immunomodulation

B. clausii has been shown to have immunomodulatory properties in preclinical studies. In a human enterocyte model of rotavirus infection, *B. clausii* strains (O/C, SIN, N/R, and T) induced the synthesis of bacteriocins, reduced enterocyte cell death, and inhibited the release of pro-inflammatory cytokines. They also increased mucin production and the synthesis of tight junction proteins, both important for the integrity of the gut mucosal barrier.[30] In addition, a small *in vivo* experiment has shown that *B. clausii* modifies the gene expression profile in the intestine in patients with mild oesophagitis, including genes involved in immunity and inflammation.[31] Finally, in an animal model of asthma, *B. clausii* reduced the numbers of eosinophils, neutrophils, and lymphocytes, and lowered IL-4 and IL-5 levels,

suggesting a potential use in reducing airway inflammation in clinical settings.[32]

Abreu explained that one potential mechanism for the immunomodulatory capacity of probiotic *B. clausii* strains could be the expression of extracellular compounds and/or immunostimulation via the cell wall. In murine cell lines, *B. clausii* MTC 8326 was shown to activate metabolic activity and innate immune responses in macrophages,[33] and *B. clausii* (O/C, N/R, SIN, and T) was also shown to stimulate the production of nitrite in peritoneal cells, IFN-γ in spleen cells, and CD4+ T-cell proliferation.[20] One route through which *B. clausii* may induce these immunomodulatory effects is through the secretion of lipoteichoic acid.[34]

Gut Homeostasis

Other studies have suggested that *B. clausii* contributes to gut homeostasis. In an *in vitro* simulation of the human gastrointestinal tract, *B. clausii* SC 109 spores (along with other probiotic bacteria and prebiotic ingredients) were shown to increase microbiome production of butyrate, and the overall diversity of gut microbiota.[35] The presence of *B. clausii* in patients with pancreatic adenocarcinoma has been associated with longer survival times,[36] and treatment with *B. clausii* UBBC07 has been shown to reduce serum urea levels in rats with acetaminophen-induced renal failure, suggesting a novel clinical use for probiotics in chronic kidney disease.[37]

Antimicrobial Properties

Abreu explained that *B. clausii* can produce antimicrobial peptides, including lantibiotics, that inhibit the growth of pathogenic bacteria *in vitro.*[20] This characteristic means that probiotics can be supportive when delivered alongside antibiotic therapy. *B. clausii* (O/C, N/R, SIN, and T) appears to be protective during *Escherichia coli* infection in mice, increasing protective mucus secretion and resulting in minimal mucosal damage and less sloughing of villus tips.[38,39]

Infection with *C. difficile* can result in symptoms ranging from diarrhoea to pseudomembranous colitis,[40] and infection with *B. cereus* can cause vomiting, diarrhoea, and haemorrhage.[41] *B. clausii* strain O/C has been shown to secrete a serine protease capable of inhibiting the cytotoxic effects of both *C. difficile* and *B. cereus in vitro.*[41]

Abreau explained that *B. clausii* has been efficaciously and safely used in humans for several decades. For example, in patients with dietary endotoxemia, believed to be caused by disruptions in gut permeability, administration of probiotic strains including *B. clausii* was associated with a 42% reduction in post-prandial serum endotoxin and reductions in pro-inflammatory markers.[42] In patients with recurrent aphthous stomatitis, a disease of the oral mucosa that results in ulcers and pain, local adjunct application of *B. clausii*, alongside glucocorticoid treatment, reduced oral pain and ulcer severity compared to glucocorticoid alone.[43]

In summary, Abreu reiterated that the physiological, antimicrobial, and immunomodulatory properties of *B. clausii* have been demonstrated both *in vitro* and *in vivo*; and antimicrobial activity against enteropathogens such as *C. difficile* and *B. cereus* has been demonstrated, providing one potential mode of action for the efficacy of this probiotic in gastrointestinal disorders. Further clinical studies using specific strains in targeted medical conditions are needed to validate these findings, and to increase the scientific credibility of *B. clausii*.

Gut Microbiota in Cardiometabolic Diseases

Karine Clément

Clément began by emphasising that there is a heavy societal burden from cardiometabolic and nutrition-related diseases and that the gut microbiota can be considered a 'super-integrator' for many of the risk factors for mortality.[44]

Obesity, the fourth highest risk factor for mortality in Western Europe,[44] is associated with altered inter-organ cross-talk involving the intestinal tract, brain, adipose tissue, muscles, and others (PRIEST 2019). In the adipose tissue, obesity is connected to perturbed endocrine secretions, immune or inflammatory imbalances, altered angiogenesis, organelle dysfunction, altered extracellular matrix, and adipocyte hypertrophy.[45-47] The development of obesity involves the pathogenic remodelling

of white adipose tissue, which may lead to the development of obesity-related cardiometabolic disease and compromised response to obesity treatment.[48] There is substantial heterogeneity in the clinical trajectory of subjects with obesity and their weight loss responses, for which gut microbiota-derived elements may be contributing factors.[49]

Clément explained that the role of the gut microbiota genomes in host biology should be considered: while it is accepted that both environmental and genetic factors play a role in the development of metabolic disease, gut microbiota may represent the missing link between them.

The key functions of the gut microbiota are in the digestion of food and the production of metabolites, the development and integrity of intestinal structure, immune system development, metabolism of toxic compounds, and synthesis of vitamins K and B.[50] However, several studies have suggested that gut microbiota also play a role in energy balance and our capacity to store fat. Clément described ground-breaking pre-clinical experiments that showed that germ-free rodents have decreased adiposity and are resistant to diet-induced weight gain, compared to conventionally raised rodents.[51] In addition, transplanting gut microbiota from mouse models of obesity into germ-free mice can partially transfer the obesity phenotype.[51] Similar experiments have been conducted to transfer microbiota from humans to mice, and these have shown that the receipt of gut microbiota from an obese human can result in increased adiposity in a mouse, even when a healthy diet is followed. In parallel, the receipt of gut microbiota from a lean individual (a twin of the obese individual) results in a lean mouse when a healthy diet is followed[52,53] (Figure 2).

Clément then discussed the importance of diversity in the gut microbiome in healthy individuals. Subjects living in westernised countries such as the USA have been shown to have a lower diversity of gut microbiota from an early age, compared to populations that are more isolated or live with an ancestral mode, such as Malawians or Amerindians.[54] Some studies have attempted to stratify individuals by their microbiotic gene richness. Across these studies, 20–30% of subjects were considered to have low gene richness, and this group was characterised by increased overall adiposity, dysmetabolism, and a more pronounced inflammatory phenotype than individuals with high gene richness.[55,56] Approximately 75% of patients with severe obesity (candidates for bariatric surgery) can be classified as having low gene richness.

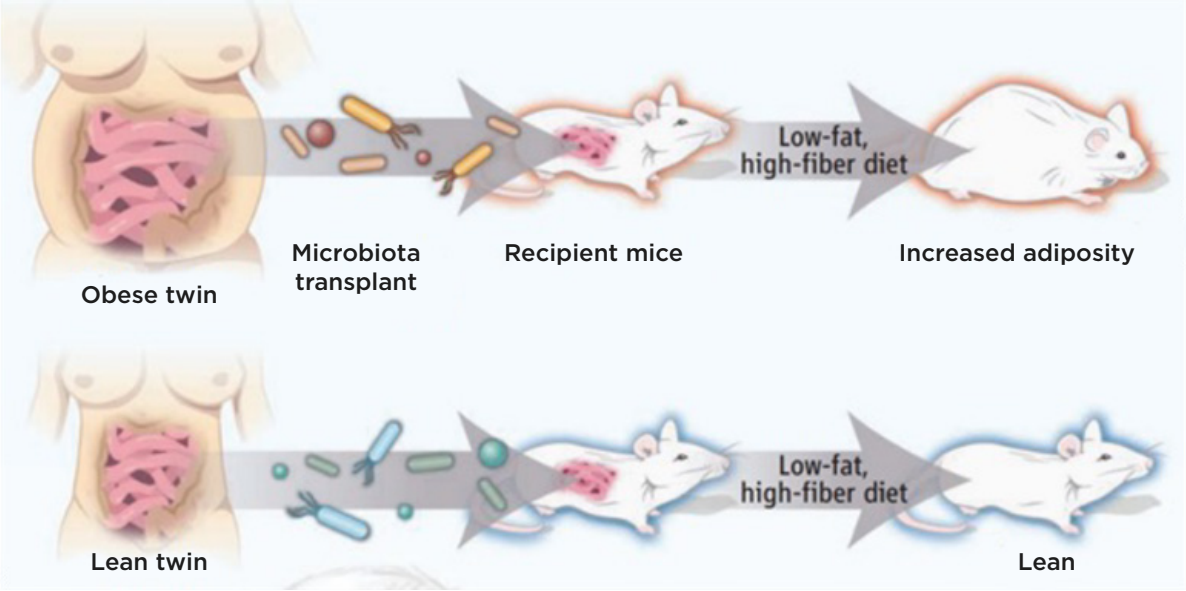


Figure 2: The protective role of gut microbiota from a lean donor in the presence of a healthy diet.

Reproduced with permission, Walker and Parkhill.[52]

This is important because a low gene count is associated with enrichment of pro-inflammatory bacteria, whereas a high gene count is associated with enrichment of anti-inflammatory bacteria.[55]

One of the important characteristics of 'healthy' gut microbiota is the production of short-chain fatty acids (SCFAs), including butyrate.[6,8,25,57] SCFAs act on enterocytes to stimulate the production of certain hormones, improving insulin sensitivity and glucose tolerance and modifying lipid metabolism.[8,25] Clément emphasised that there is considerable research effort focused on understanding the imbalance between the gut microbiota in healthy individuals and those with disease. Gut microbiota may also contribute to the health of the intestinal barrier in metabolic diseases.[58] For example, studies have shown that modification of the gut microbiota affects the thickness of the mucus barrier.[58]

The effects of gut microbiota on the host can be classified as metabolism-independent pathways, driven by components of the bacterial membrane such as lipopolysaccharide or peptidoglycan and impacting low-grade inflammation processes or modifying host biology; or metabolism-dependent pathways driven by microbial metabolites such as imidazole propionate, SCFAs, secondary bile acids, or trimethylamine. [59,60]

Clément described several studies that have attempted to stratify gut microbiomes into groups based on their genome. In a European study, Arumugam et al., described three distinct clusters of microbiomes, termed enterotypes, each characterised by a dominant gut microbial species: Type 1, enriched in *Bacteroides*; Type 2, enriched in *Prevotella;* and Type 3, enriched in *Ruminococcus.*[61] Subsequent studies have identified a subset of the Type 2 microbiome with a low proportion of *Faecalibacterium* and low microbial cell density, named Bact2, which is more prevalent in patients with inflammatory bowel disease versus the general population (78% versus 13%, respectively).[62] The prevalence of Bact2 also correlates with higher BMI and with low-grade systemic inflammation in the MetaCardis European cohort.[63]

Clément explained that interventions to increase microbial diversity, increase beneficial microbes, and change metabolite concentrations are intended to improve metabolism and the immune response, potentially reducing the burden of complications. Potential mechanisms to modify the gut microbiome include dietary changes, selective enrichment of gut bacteria, faecal transplant, and bariatric surgery.

One example of such an intervention is diet-induced weight loss in patients with obesity or overweight, which improved gut microbiotic diversity and clinical phenotypes in patients with a low microbial gene count at baseline.[56] Bariatric surgery also appears to increase microbial gene richness one-year post-surgery.[63] Administration of *Akkermansia muciniphila* to mouse models of obesity or Type 2 diabetes resulted in a reduction in fat mass, insulin resistance, and low-grade inflammation,[64] and *A. muciniphila* is associated with healthier metabolic status and greater insulin sensitivity in human subjects with obesity or overweight.[65] Finally, a study of faecal transfer from healthy individuals to patients with obesity and metabolic syndrome showed an improvement in insulin sensitivity, however, the effect was transient and mainly observed in patients with a low gut microbiota diversity at baseline.[66]

Clément concluded that there is evidence for a dysbiosis of gut microbiota in metabolic diseases, and that a personalised approach to the gut microbiome may be the best way to leverage this association. However, she stressed that further research is needed as the links between the changes in the gut microbiota and the expected clinical effects have yet to be fully elucidated.

Summary

In summary, the evidence to date supports the hypothesis that both probiotics and the gut microbiome have an impact on the health of humans and other animals. However, though potential mechanisms of action have been suggested experimentally, further research including well-designed trials is needed to fully understand how probiotics manipulate the gut microbiota to benefit the host.

MAT-GLB-2104926

References

1. Laukens D et al. Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. FEMS Microbiol Rev. 2016;40(1):117–132.
2. Merenstein DJ et al. Probiotics as a Tx resource in primary care. J Fam Pract. 2020;69(3):E1–E10.
3. Wang B et al. The human microbiota in health and disease. Engineering. 2017;3(1):71–82.
4. McBurney MI et al. Establishing what constitutes a healthy human gut microbiome: state of the science, regulatory considerations, and future directions. J Nutr. 2019;149(11):1882–95.
5. Hill C et al. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol. 2014;11(8):506–14.
6. Gibson GR et al. Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. Nat Rev Gastroenterol Hepatol. 2017;14(8):491–502.
7. Swanson KS et al. The International Scientific Association for Probiotics and Prebi-otics (ISAPP) consensus statement on the definition and scope of synbiotics. Nat Rev Gastroenterol Hepatol. 2020;17(11):687–701.
8. Salminen S et al. The International Scientific Association of Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of postbiotics. Nat Rev Gastroenterol Hepatol. 2021;18(9):649–67.
9. McFarland LV et al. Strain-specificity and disease-specificity of probiotic efficacy: a systematic review and meta-analysis. Front Med (Lausanne). 2018;5:124.
10. International Scientific Association for Probiotics and Prebiotics. ISAPP position statement on minimum criteria for harmonizing global regulatory approaches for probiotics in foods and supplements. 2018. Available at: https://isappscience.org/minimum-criteria-probiotics. Last accessed: October 2021.
11. World Health Organisation. Probiotics in food Health and nutritional properties and guidelines for evaluation. 2006. Available at: http://www.fao.org/3/a0512e/a0512e.pdf. Last accessed: October 2021.
12. Dailey Z et al. Retail refrigerated probiotic foods and their association with evidence of health benefits. Benef Microbes. 2020;11(2):131–3.
13. Merenstein D et al. More information needed on probiotic supplement product labels. J Gen Intern Med. 2019;34(12):2735–7.
14. Kristensen NB et al. Alterations in fecal microbiota composition by probiotic supplementation in healthy adults: a systematic review of randomized controlled trials. Genome Medicine. 2016;8(1):52.
15. Khurana H et al. Genomic insights into the phylogeny of Bacillus strains and elucidation of their secondary metabolic potential. Genomics. 2020;112(5):3191–3200.
16. Elshaghabee FMF et al. Bacillus as potential probiotics: status, concerns, and future perspectives. Front Microbiol. 2017;8:1490.
17. Cutting SM et al. Bacterial spore-formers: friends and foes. FEMS Microbiol Lett. 2014;358(2):107–9.
18. Bacillus clausii Enterogermina.png. Wikimedia Commons. 2018. Available at: https://commons.wikimedia.org/wiki/File:Bacillus_clausii_Enterogermina.png. Last accessed: October 2021
19. Khatri I et al. Composite genome sequence of Bacillus clausii, a probiotic commer-cially available as Enterogermina®, and insights into its probiotic properties. BMC Microbiology. 2019;19(1):307.
20. Urdaci MC et al. Bacillus clausii probiotic strains antimicrobial and immunomodulatory activities. J Clin Gastroenterol. 2004;38(6 Suppl):S86–S90.
21. Kolacek S et al. Commercial probiotic products: a call for improved quality control. a position paper by the ESPGHAN working group for probiotics and prebiotics. J Pediatr Gastroenterol Nutr. 2017;65(1):117–24.
22. Senesi S et al. Molecular characterization and identification of Bacillus clausii strains marketed for use in oral bacteriotherapy. Appl Environ Microbiol. 2001;67(2):834–9.
23. Vecchione A et al. Compositional quality and potential gastrointestinal behaviour of probiotic products commercialized in Italy. Front Med (Lausanne). 2018;5:59.
24. Lee SH et al. Isolation and physiological characterization of Bacillus clausii SKAL-16i from wastewater. J Microbiol Biotechnol. 2008;18(12):1908–14.
25. Cantu-Jungles TM et al. Potential of prebiotic butyrogenic fibers in parkinson's dis-ease. Front Neurol. 2019;10:663.
26. Lakshmi SG et al. Safety assesment of Bacillus clausii UBBC07, a spore forming probiotic. Toxicol Rep. 2017;4:61-71.
27. Salvetti S et al. Rapid determination of vitamin B2 secretion by bacteria growing on solid media. J Appl Microbiol. 2003;95(6):1255–60.
28. Ahire JJ et al. Survival and Germination of Bacillus clausii UBBC07 Spores in in vitro Human Gastrointestinal Tract Simulation Model and Evaluation of Clausin Produc-tion. Front Microbiol. 2020;11:1010.
29. Bouhss A et al. Specific interactions of clausin, a new lantibiotic, with lipid precursors of the bacterial cell wall. Biophys J. 2009;97(5):1390–7.
30. Paparo L et al. Protective action of Bacillus clausii probiotic strains in an in vitro model of Rotavirus infection. Sci Rep. 2020;10(1):12636.
31. Di Caro S et al. Bacillus clausii effect on gene expression pattern in small bowel mucosa using DNA microarray analysis. Eur J Gastroenterol Hepatol. 2005;17(9):951–60.
32. Park H et al. Bacillus clausii, a foreshore-derived probiotic, attenuates allergic airway inflammation through downregulation of hypoxia signaling. J Rhinol. 2020;27(2):108–116.
33. Pradhan B et al. Comparative analysis of the effects of two probiotic bacterial strains on metabolism and innate immunity in the raw 264.7 murine macrophage cell line. Probiotics Antimicrob Proteins. 2016;8(2):73-84.
34. Villéger R et al. Characterization of lipoteichoic acid structures from three probiotic Bacillus strains: involvement of D-alanine in their biological activity. Antonie Van Leeuwenhoek. 2014;106(4):693–706.
35. Duysburgh C et al. A synbiotic concept containing spore-forming Bacillus strains and a prebiotic fiber blend consistently enhanced metabolic activity by modulation of the gut microbiome in vitro. Int J Pharm X. 2019;1:100021.
36. Riquelme E et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. Cell. 2019;178(4):795–806.
37. Patel C et al. Therapeutic prospective of a spore-forming probiotic-Bacillus clausii UBBC07 against acetaminophen-induced uremia in rats. Probiotics Antimicrob Pro-teins. 2020;12(1):253–8.
38. Yu MG et al. Histomorphologic effects of Bacillus clausii spores in enteropathogenic E. coli O127:H21-infected mice: A Pilot Study. Phillipine J Int Med. 2016;54:1–7.
39. De Castro JA et al. Bacillus clausii as adjunctive treatment for acute community-acquired diarrhea among Filipino children: a large-scale, multicenter, open-label study (CODDLE). Tropical Dis Travel Med Vacc. 2019;5:14.

40. Mills JP et al. Probiotics for prevention of clostridium difficile infection. Curr Opin Gastroenterol. 2018;34(1):3–10.
41. Ripert G et al. Secreted compounds of the probiotic Bacillus clausii strain O/C inhibit the cytotoxic effects induced by Clostridium difficile and Bacillus cereus toxins. Antimicrob Agent Chemother. 2016;60(6):3445–54.
42. McFarlin BK et al. Reversing meal-associated gastrointestinal gut permeability issues: potential treatment target for spore-based probiotics? Am J Gastroenterol. 2017;112:S658–59.
43. Cheng B et al. The efficacy of probiotics in management of recurrent aphthous stomatitis: a systematic review and meta-analysis. Sci Rep. 2020;10(1):21181.
44. IHME, Global Burden of Disease (2017). Number of deaths by risk factor, Western Europe. 2017. Available at: https://ourworldindata.org/grapher/number-of-deaths-by-risk-factor?country=-Western+Europe. Last accessed: October 2021.
45. Kasinska MA et al. Epigenetic modifications in adipose tissue – relation to obesity and diabetes. Arch Med Sci. 2016;12(6):1293–1301.
46. Nijhawans P et al. Angiogenesis in obesity. Biomed Pharmacother. 2020;126:110103.
47. Xiao Y et al. Chronic stress, epigenetics, and adipose tissue metabolism in the obese state. Nutr Metab (Lond). 2020:17:88.
48. Henegar C et al. Adipose tissue transcriptomic signature highlights the pathological relevance of extracellular matrix in human obesity. Genome Biol. 2009:9(1):R14.
49. Hung TKW et al. Understanding the heterogeneity of obesity and the relationship to the brain-gut axis. Nutrients. 2020;12(12):3701.
50. Fouhy F et al. Composition of the early intestinal microbiota: knowledge, knowledge gaps and the use of high-throughput sequencing to address these gaps. Gut Microbes. 2012;3(3):203-20.
51. Xiao H et al. The role of the gut microbiome in energy balance with a focus on the gut-adipose tissue axis. Front Genet. 2020;11:297.
52. Ridaura VK et al. Cultured gut microbiota from twins discordant for obesity modulate adiposity and metabolic phenotypes in mice. Science. 2013;341(6150):1241214.
53. Walker AW and Parkhill J. Fighting obesity with bacteria. Science. 2013:341(6150):1069–70.
54. Yatsunenko T et al. Human gut microbiome viewed across age and geography. Na-ture. 2012;486(7402):222–7.
55. Le Chatelier E et al. Richness of human gut microbiome correlates with metabolic markers. Nature. 500(7464):541–546.
56. Cotillard A et al. Dietary intervention impact on gut microbial gene richness. Nature. 2013;500(7464):585–8.
57. Montassier E et al. Chemotherapy-driven dysbiosis in the intestinal microbiome. Aliment Pharmacol Ther. 2015;42(5):515–28.
58. Cani PD et al. Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. Gut. 2009:58(8):1091–103.
59. Brown JM et al. The gut microbial endocrine organ: bacterially-derived signals driving cardiometabolic diseases. Annu Rev Med. 2015;66:343–59.
60. Koh A et al. Microbially produced imidazole propionate impairs insulin signaling through mTORC1. Cell. 2018;175(4):947–61.e17.
61. Arumugam M et al. Enterotypes of the human gut microbiome. Nature. 2011;473(7346):174–80.
62. Vieira-Silva S et al. Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. Nature. 2020;581(7808):310–5.
63. Aron-Wisnewsky J et al. Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. Gut. 2019;68(1):70–82.
64. Everard A et al. Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. Proc Natl Acad Sci USA. 2013;110(22):9066-71.
65. Dao MC et al. Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. Gut. 2016;65(3):426–36.
66. Kootte RS et al. Improvement of insulin sensitivity after lean donor feces in metabolic syndrome is driven by baseline intestinal microbiota composition. Cell Metab. 2017;26(4):611–9.e6.