



Prognostic value of macrophage polarization markers in epithelial neoplasms and melanoma. A systematic review and meta-analysis

Álvaro López-Janeiro¹ · Carlos Padilla-Ansala¹ · Carlos E. de Andrea^{2,3,4,5} · David Hardisson^{1,6,7} · Ignacio Melero^{4,5,8,9}

Received: 30 January 2020 / Revised: 14 March 2020 / Accepted: 17 March 2020 / Published online: 14 April 2020
© The Author(s), under exclusive licence to United States & Canadian Academy of Pathology 2020

Abstract

Macrophage polarization is relevant for tumor biology. M2 polarized macrophages favor tumor growth and survival, while M1 macrophages support tumor destruction and antigen presentation. Markers identifying M1/M2 polarization are a subject of debate. We conducted a systematic review and meta-analysis to investigate the association of proposed macrophage markers with prognosis across epithelial tumors and melanoma. The Medline search engine was used and 195 articles were recovered for full review. Only articles which measured markers using immunohistochemistry or immunofluorescence and had overall survival (OS) as the primary endpoint were included. One hundred and thirteen articles were finally accepted for analysis. CD68 was associated with worse survival across tumors (hazard ratio (HR) = 1.24, 95% CI = 1.11–1.37). Tumor anatomical location influenced this association. Colorectal tumors showed an inverse association between CD68 and OS in contrast to the rest of cancer types (HR = 0.56 vs. 1.34). The approach taken to measure CD68 had an impact on prognosis; when macrophages were measured at the tumor invasion front prognosis was more favorable than when they were measured intratumorally (HR = 0.94 vs. 1.4). CD163, CD204, and CD206 showed a robust association with worse OS (HR = 1.63, 1.95, 1.65, respectively). Tumors arising in the lung and the liver showed a weaker association between CD163 and OS as compared with other locations ($\beta = -0.5401$ for the lung and -0.5940 for the liver compared with other anatomical locations). The counting strategy also had an impact on CD163 association with OS, with hot-spot counting having higher HRs compared with averaging macrophage counts across spots or absolute cell counting ($\beta = -0.4678$). In conclusion, proposed M2 markers are associated with worse survival across epithelial tumors and melanoma. The anatomical origin of tumors influences this association. The compartment where the macrophages were scored and counting strategy influenced the association with OS of CD68 and CD163, respectively.

Supplementary information The online version of this article (<https://doi.org/10.1038/s41379-020-0534-z>) contains supplementary material, which is available to authorized users.

✉ Álvaro López-Janeiro
alopez.44@alumni.unav.es

¹ Department of Pathology, Hospital Universitario La Paz—IdiPaz, Madrid, Spain

² Department of Pathology, Clínica Universidad de Navarra, University of Navarra, Pamplona, Spain

³ Department of Anatomy, Physiology and Pathology, University of Navarra, Pamplona, Spain

⁴ Navarra Institute for Health Research (IDISNA), Pamplona, Spain

Introduction

Macrophages play a key role in tissue repair and immune response to pathogens. The wide variety of functions they perform is due to their ability to adapt and adjust to the

⁵ Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain

⁶ Molecular Pathology and Therapeutic Targets Group, La Paz University Hospital—IdiPaz, Madrid, Spain

⁷ Faculty of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

⁸ Department of Immunology and Immunotherapy, Clínica Universidad de Navarra, Pamplona, Spain

⁹ Program of Immunology and Immunotherapy, CIMA Universidad de Navarra, Pamplona, Spain

requirements of the tissue microenvironment [1]. Two main functional subtypes of macrophages have been described: M1 and M2. M1-macrophage functions include antigen presentation and intracellular pathogen or tumor cell destruction among others. M2 macrophages promote extracellular matrix remodeling, angiogenesis and express immunomodulatory features. In tumors, infiltrating macrophages, known as tumor-associated macrophages (TAM), are usually educated by environmental factors to differentiate into M2 state. The M1- and M2-subtypes represent a spectrum of functionally diversified immune cells. Macrophage polarization, the process that skews macrophages toward specific functions, is associated with tumor cell biology. M2-polarized macrophages are associated with tumor growth and progression, whereas M1 macrophages show pro-inflammatory and tumoricidal properties. However, M1 and M2 macrophages are difficult to characterize due to the great variety of functional phenotypes and absence of specific markers [2, 3]. While CD68 is considered the gold standard marker for human macrophages, there is still little consensus about which surface markers can be used to identify macrophage functional states. Since the discovery of macrophage plasticity, several markers that identify macrophage polarization subtypes have been described [4]. Due to the increased scavenging capabilities of M2 polarized macrophages, the scavenger receptors CD163 (Hemoglobin-Haptoglobin SCR), CD204 (Scavenger Receptor A), and CD206 (Mannose Receptor C type 1) have been proposed as markers of M2-skewed macrophages [5]. On the other hand, as M1 macrophages show increased antigen presentation, lymphocyte co-stimulation, and killing properties, markers related to these processes, including HLA-DR, CD83, CD80, CD40, and inducible nitric oxide synthase, have been proposed to identify this cell population.

Tumor cells can interact with macrophages and induce them to acquire a more M2-like phenotype, tampering with their tumoricidal capability and taking advantage of their tissue remodeling and pro-angiogenic functions to promote tumor growth [6, 7]. The prognostic significance of TAM has been shown in several studies. In general, TAMs are predictors of worse outcome [5, 6]. Furthermore, the prognostic significance of TAM has shown to be organ related [7, 8].

Here, we present a systematic review and meta-analysis of studies assessing the impact of different TAM markers on disease outcome across solid epithelial tumors and melanoma. Our aim is to study the prognostic significance of each individual marker and identify variables that modify this association.

Material and methods

The Pubmed search engine was used to conduct the study. Two queries with the terms “macrophages” (major MESH

term) AND “prognosis” AND “cancer” plus “macrophages” (major MESH term) AND “prognosis” AND “tumor” were conducted. Results were filtered by date of publication including only those articles published between January 1, 2003 and July 31, 2018 and those written in English. After removing duplicates, abstracts were screened and articles not fulfilling the inclusion criteria were excluded. The full text of the remaining articles was reviewed. In addition, studies included in any of the published systematic reviews or meta-analyses identified in our search were also included for review in the present work. Inclusion criteria included the following: (1) prospective and retrospective cohort and case control studies, including patients with any type of epithelial tumor or melanoma at any stage of diagnosis with clinical follow-up were considered for review. (2) Patients included in the study should not have received anti-PD1/PDL1 or anti-CTLA-4 treatment at any point in their disease. (3) Only macrophage surface markers and non-secreted factors were allowed in the review. (4) Macrophage markers had to be measured using immunohistochemistry (IHC) or immunofluorescence (IF) and cut-off values with the number of patients allocated to each category for group comparison had to be specified. (5) The measurement had to be taken at the primary tumor site. Articles measuring macrophages in metastases or local relapses were excluded as were case reports, animal model-based studies, and literature reviews. Studies not measuring macrophage markers with IHC or IF or measuring extracellular-secreted products (IE: interleukins) were also excluded from the analysis. The main clinical outcome studied in this systematic review was overall survival (OS), although data regarding disease-free survival (DFS) or progression-free survival (PFS) were also gathered if presented.

Two independent researchers gathered the information from the studies. Variables collected were grouped into four categories: bibliographic data (including year of publication and journal), characteristics of the sample included in the study (tumor histologic type, anatomic location, stage at diagnosis, treatment received, mean/median age, age range, and male/female ratio), macrophage marker assessment specifications (IHC- or IF-based, clone used for determination, compartment where the measurement was taken (i.e., intra-tumoral or stromal), image analysis assisted or naked eye evaluation, quantitative or qualitative assessment and if assessment was done blinded to clinical outcome), and outcome-related variables (follow-up of the sample, outcome type, events recorded at the end of follow-up, cut-off percentile for high/low group splitting, association measure type, adjusted or unadjusted measure provided, association measure with 95% confidence interval (CI) and adjustment variables). Hazard ratio (HR) was the summary measure adopted in this systematic review. Those articles

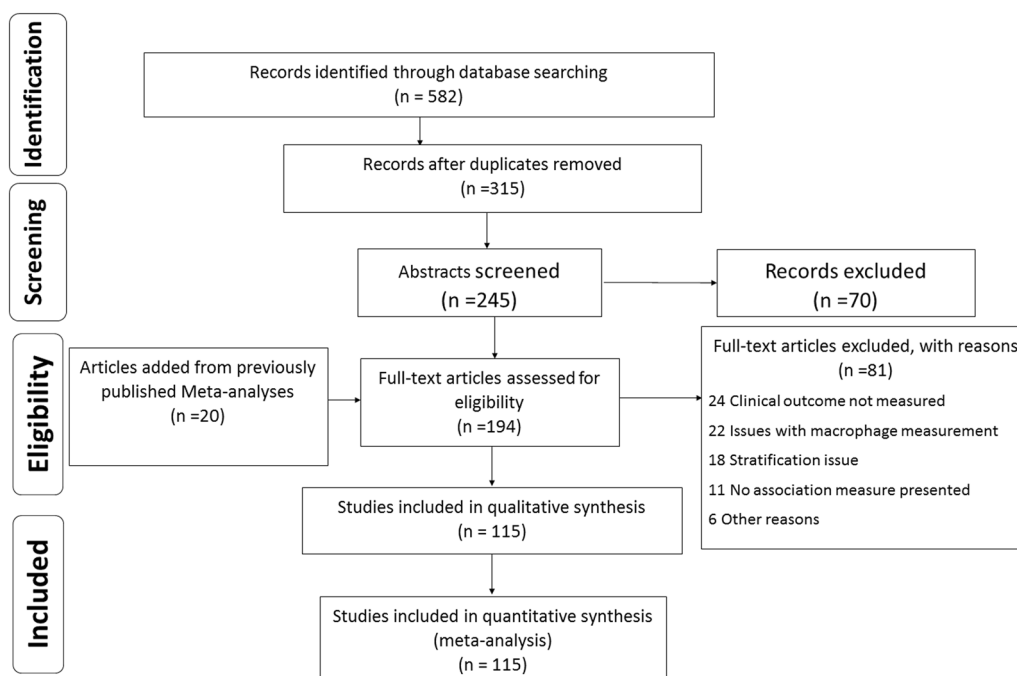


Fig. 1 PRISMA flowchart. Flowchart outlining the article selection process.

that did not provide HR and CI but presented a Kaplan–Meier (KM) plot and a p value derived from log-rank analysis were also accepted. HR and standard error were calculated from the KM plot and the p value as previously described [9]. The hazard ratio was calculated taking the lower macrophage containing group as a reference. If the highest macrophage containing group was used as reference in the article, then the association measure and CI were inverted.

All statistical analyses were conducted using R (3.5.1) and the Rcmdr package [10]. The METAFOR package [11] was used to conduct the meta-analysis and analyze moderator variables. Random effect models were used to calculate aggregated results, and mixed-effects models were calculated to demonstrate the impact of different variables on the aggregated measure. The between-study variance (τ^2) was calculated using DerSimonian-Laird method. The Knapp and Hartung adjustment was used to penalize p values and those whose omnibus test resulted in $p < 0.05$ were further screened using a permutation test (1000 iterations or exact permutation test if there were less than 1000 possible permutations). Before analysis data were inspected and extreme outliers (a reported HR deviating more than 3 standard deviations from mean) were removed from the analysis. Funnel and radial plots were visually inspected to assess publication bias. If publication bias was observed, the trim and fill data augmentation technique was used to calculate non-biased aggregated estimations [12]. Plots were constructed using forestplot and ggplot2 packages [13, 14].

Results

The sum of both search queries yielded 582 results. Three hundred and fifteen articles remained after removal of duplicates. Previously published systematic reviews and meta-analyses yielded 20 additional articles that were not present in our Medline search. After abstract screening, 194 full text articles were considered for review. Eighty-one articles were rejected due to the following reasons: 24 did not present a correctly assessed clinical outcome, 22 did not meet the requirements related to macrophage assessment, 18 did not state the cut-off point to stratify the sample into groups in order to calculate an association measure, 11 did not present an association measure, and 6 were excluded due to miscellaneous reasons (duplicate studies and severe methodological shortcomings) (the PRISMA [15] article selection flowchart is shown in Fig. 1).

Of the 113 articles that were accepted for qualitative analysis, 2 presented two independent cohorts of patients with independent measures yielding 115 articles that were effectively reviewed. Information regarding tumor anatomical location and patient demographics is shown in Fig. 2.

CD68

Seventy-seven raw and 34 adjusted measurements analyzed the association of CD68 with OS. These were provided by 59 articles. Information regarding the populations studied and the material and methods used by each article is summarized in Supplementary Table 1. Aggregated analysis

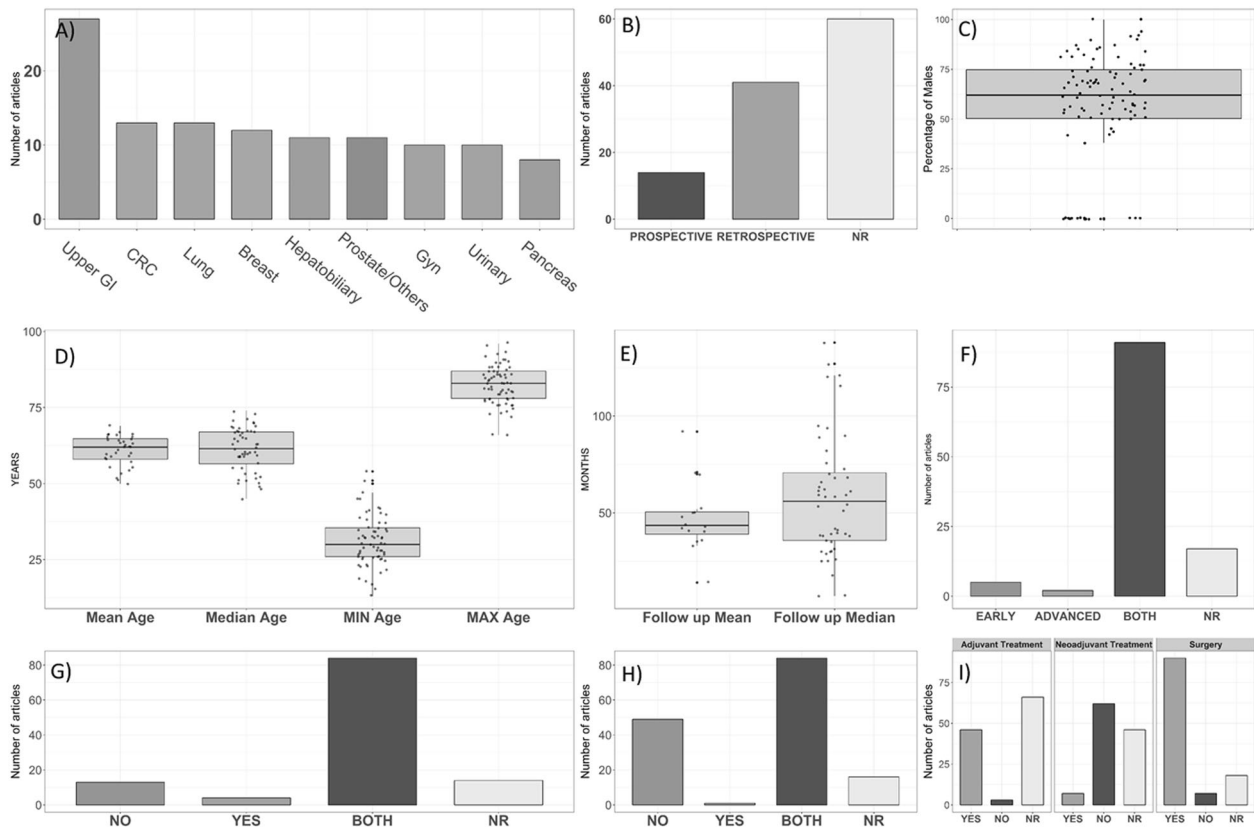


Fig. 2 Summary of articles reviewed in the meta-analysis. **a** Location of tumors studied in articles. Upper GI includes tumors at esophageal, gastric, and ampullary locations. Prostate/others include prostatic, ocular, and skin tumors. Gynecological cancers include cervical, uterine, and ovarian cancer. Urinary locations include kidney and bladder tumors. **b** Design of studies included in the analysis. NR not reported. **c** Mean percentage of males included in each article. Data not reported in 13 out of 115 (11%) articles. **d** Mean, median, minimum, and maximum age of patients included in each article. Sixty-three out of 115 articles (55%) did not report mean age of patients. Eighty-one out of 115 articles (70%) did not report the median age of patients. Forty-eight out of 115 (42%) did not report the minimum or maximum

showed an association of CD68 with OS (aggregated HR = 1.24, 95% CI = 1.11–1.37) for raw measures. Aggregated association of adjusted measures with OS was less robust (HR = 1.01, 95% CI = 0.92–1.1). I^2 indices of residual heterogeneity for aggregated raw and adjusted measures were 82% and 73%, respectively (Fig. 3). Results regarding the association of CD68 with other endpoints (DFS or PFS) are summarized in Supplementary Table 2.

The compartment where the macrophages were measured and the anatomical location of tumors profoundly modified the raw CD68 association with OS (Fig. 3). While measurements taken in the tumor compartment yielded an aggregated HR of 1.4, the aggregated HR for measures taken at the invasive front was 0.94 ($\beta = -0.4$, 95% CI = -0.77 to -0.02 permutated p value = 0.024). Interestingly, CD68 measurements from tumors located in the lower gastro-intestinal region (colon and rectum) showed an

age of patients. **e** Mean and median follow-up of each article. Ninety-nine and 67 out of 115 articles (86 and 58%) did not report mean and median follow-up, respectively. **f** Local involvement of tumors of the patients studied in each article at diagnosis. NR not reported. **g** Lymph node involvement of the patients studied in each article at diagnosis. NR not reported. **h** Presence of metastasis at diagnosis in the patients studied in each article. NR not reported. **i** Treatment received by patients included in each article. Any form of radiotherapy or chemotherapy is considered as neoadjuvant or adjuvant therapy. NR not reported, Upper GI upper gastro intestinal, CRC colorectal carcinoma, Gyn: gynecological tumors.

inverse relationship with OS (HR = 0.56, 95% CI = 0.38–0.83), as compared with any other location (HR = 1.34, 95% CI = 1.18–1.53) ($\beta = 0.8808$, 95% CI = 0.4658–1.2958 permutated p value = 0.001).

CD163

Thirty-three studies provided a total of 46 raw and 29 adjusted measures analyzing the relationship between CD163 and OS. Supplementary Table 3 outlines the main characteristics of these articles. The aggregated HR of CD163 with OS was 1.63 (95% CI 1.42–1.86) for raw measures and 1.16 (95% CI 1.1–1.23) for adjusted measures (Fig. 4). I^2 scores for residual heterogeneity were 73.42% for raw measures and 91.87% for adjusted measures.

Analysis of variables possibly influencing the relationship between CD163 and OS showed that the anatomical

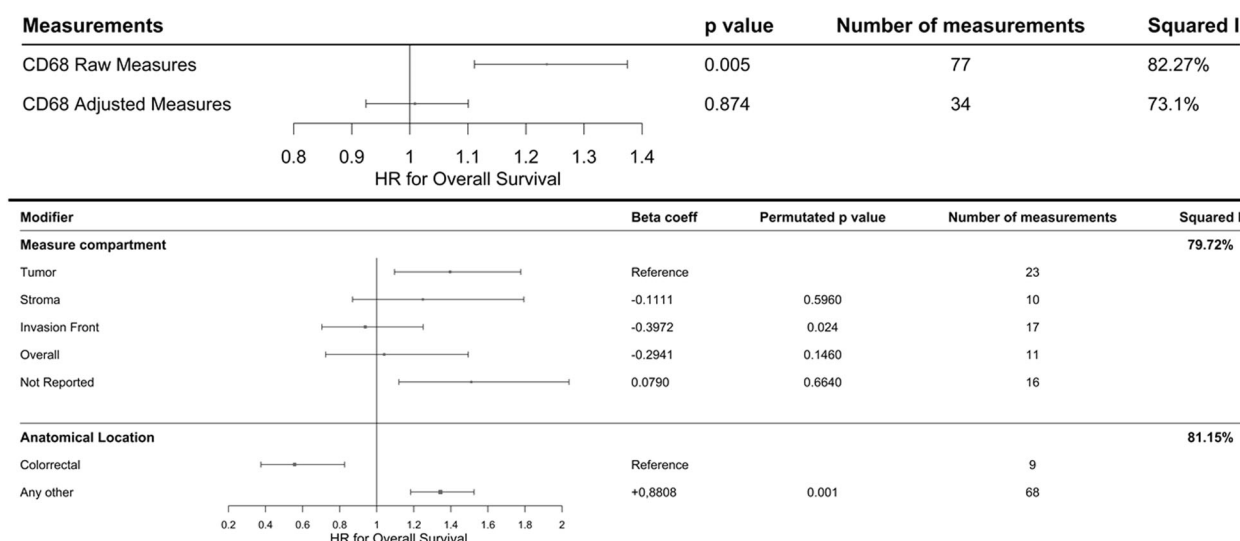


Fig. 3 Upper panel: aggregated hazard ratio and 95% CI for the association of OS with CD68 expression. The size of the square is inversely proportional to the size of the CI. Lower panel: aggregated

hazard ratio and 95% CI for the variables influencing the association of CD68 with OS for raw measures. β coefficients are shown. The size of the square is inversely proportional to the size of the CI.

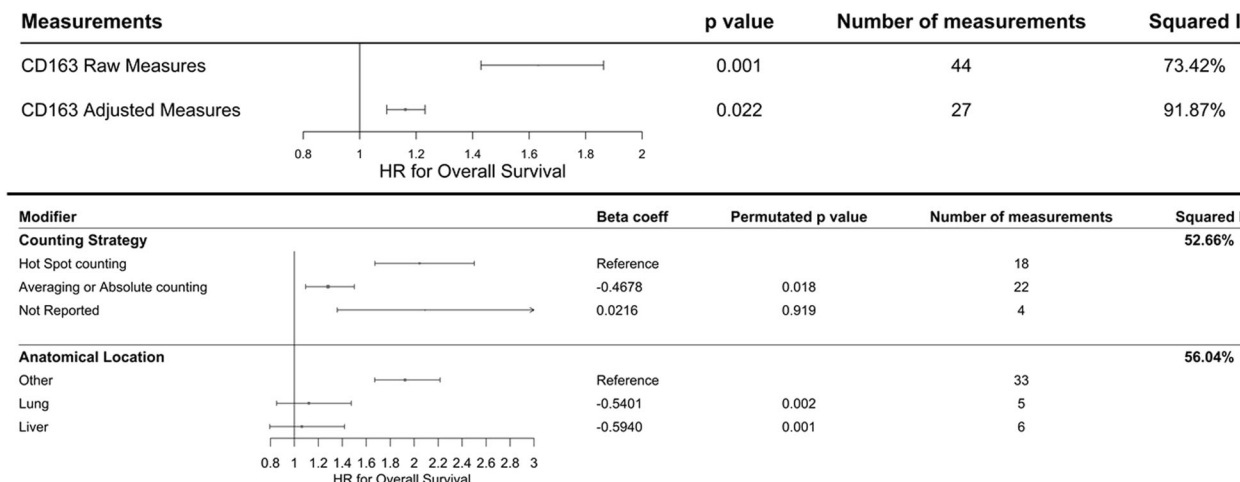


Fig. 4 Upper panel: aggregated hazard ratio and 95% CI for the association of OS with CD163 expression. The size of the square is inversely proportional to the size of the CI. Lower panel: aggregated

hazard ratio and 95% CI for the variables influencing the association of CD163 with OS for raw measures. β coefficients are shown. The size of the square is inversely proportional to the size of the CI.

location and the counting strategy were important moderators. Tumors located in the lung and liver showed a weaker relationship between CD163 and OS, compared with other tumor locations ($\beta = -0.5401$ for the lung and -0.5940 for the liver, 95% CI = -0.8484 to -0.2318 for the lung and -0.9155 to -0.2724 for the liver, permutated p value = 0.002 for the lung and 0.001 for the liver. Models built with raw measures) (Fig. 4). Articles focusing on lung cancer included patients with all histologic types of non-small cell lung cancer, while the liver cancer articles mostly included patients with hepatocellular carcinoma (HCC) with a single article studying patients with cholangiocarcinoma. These findings held true for the adjusted measurements ($\beta = -1.2197$ for the lung and -0.7912 for the liver, 95%

CI = -0.4787 to -1.1036 for the lung and -0.1386 to -2.3007 for the liver, permutated p value = 0.083 for the lung and 0.088 for the liver). On the other hand, hot-spot counting had a robust impact on the association between CD163 and OS when compared with other counting techniques such as averaging macrophage counts across several spots or absolute macrophage number counting ($\beta = -0.4678$, 95% CI = -0.7230 to -0.2125 , permutated p value = 0.001 for model built with raw measures). Similar results were found when analyzing adjusted measures ($\beta = -1.0981$, 95% CI = -1.4671 to -0.7291 , permutated p value = 0.018). Data regarding the association of CD163 with other outcomes can be found in Supplementary Table 4.

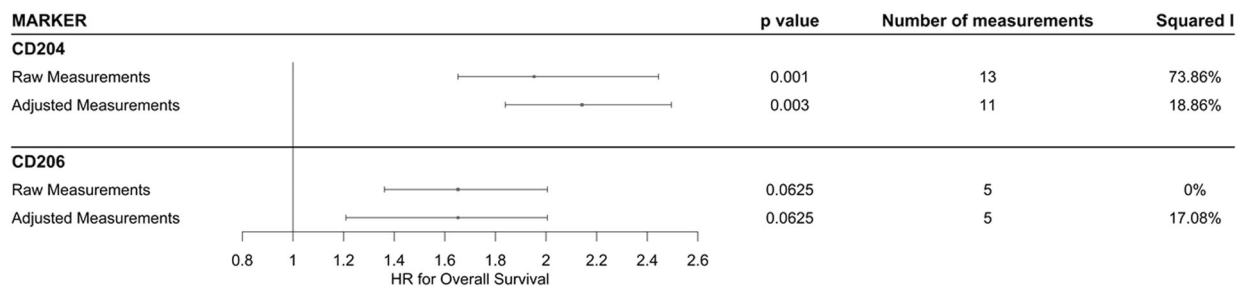


Fig. 5 Aggregated hazard ratio and 95% CI for the association of OS with CD204 and CD206 expression. The size of the square is inversely proportional to the size of the CI.

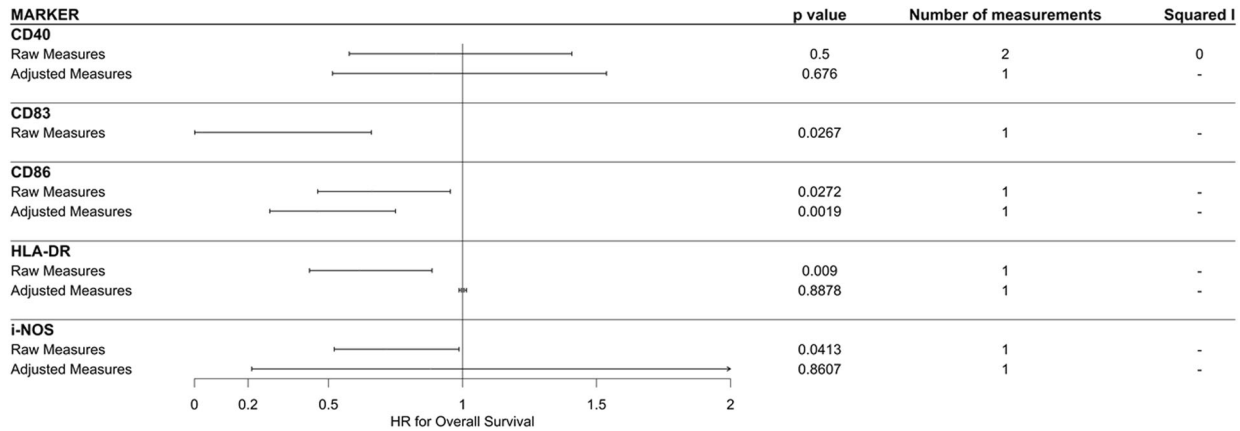


Fig. 6 Aggregated hazard ratio and 95% CI for the association of OS with the expression of several markers. The size of the square is inversely proportional to the size of the CI. iNOS inducible nitric oxide synthase.

CD204 and CD206

Thirteen and ten studies measured the association between OS and CD204 and CD206, respectively. For CD204, 13 raw and 11 adjusted measures were retrieved. A description of the articles included in the analysis can be found in Supplementary Tables 5 and 6. The aggregated HR for raw measures was robustly associated with OS (HR = 1.95, 95% CI = 1.56–2.44, *p* value = 0.001), as was the aggregated HR for adjusted measures (HR = 2.14, 95% CI = 1.84–2.5, *p* value = 0.003) (Fig. 5). CD206 results showed a similar pattern, yielding an aggregated HR of 1.65 (95% CI = 1.36–2) for raw measures and 1.57 (95% CI = 1.21–2.03) for adjusted measures.

Data regarding the influence of the anatomical location and macrophage counting strategy were non-consistent between CD204 and CD206 (Supplementary Table 7 and Supplementary Fig. 1).

M1 and other markers

Several M1 like markers were retrieved for aggregated analysis. For most of the markers, only isolated reports have been published. Supplementary Table 8 summarizes the

main characteristics of the articles studying these markers. Most showed a favorable association with OS but with variable degrees of robustness (Fig. 6).

Several articles studying other macrophage markers were also identified and included for analysis. Most of these markers have only been measured in single tumor types. The main characteristics of these papers and their results are shown in Supplementary Table 9 and Supplementary Fig. 2.

Publication bias

Slight publication bias was suspected for CD68, CD163, and CD204 based on the funnel and radial plots. For all three markers, results were biased toward an overestimated aggregated measure. The trim and fill method yielded a more conservative but still robust association with survival for the aggregated HR of CD68 (1.19 95% CI = 1.072–1.33), CD163 (1.39 95% CI = 1.22–1.58), and CD204 (1.7 95% CI = 1.36–2.11) (Supplementary Fig. 3). To obtain these estimations, the trim and fill method calculated 3, 13, and 5 additional measures for CD68, CD163, and CD204, respectively. Funnel plots generated from the analysis are shown in Supplementary Fig. 4.

Discussion

Our work reviews the association of the main single markers used to quantify macrophages with clinical endpoints across epithelial tumors and melanoma in patients who have not been treated with checkpoint inhibitors. To our knowledge this is the first systematic review of the clinical impact of macrophage markers other than CD68.

CD68 as well as the proposed M2 markers (CD163, CD204, and CD206) showed a robust association with worse clinical outcome. Importantly, while tumor location did not greatly influence the association of M2 markers with clinical outcome, it did alter the association for the general macrophage marker CD68, which is consistent with the literature [8]. Colorectal tumors harboring increased CD68 macrophage populations showed better prognosis. These findings may be related to the diverse tumor micro-environment found in colorectal tumors. The compartment where CD68 was measured also had an impact on its association with survival. Intra-tumoral CD68-positive macrophages were associated with worse outcome, while CD68-positive macrophages in the invasive front correlated with better survival. This finding possibly reflects different macrophage polarization processes occurring in these two different tumor compartments or the chemokine milieu of the malignant tissue. Interestingly, no other demographic variable (including male to female ratio, age, stage at diagnosis, or treatment received) or macrophage identification/counting method (IHC vs. IF, clone used, image analysis-based vs. naked eye counting, whole slide assays vs. TMA) influenced the association between CD68 and outcome. While aggregated raw measures for CD68 association with OS yielded a robust result, the aggregated measures for adjusted measures were less consistent, probably reflecting CD68 correlation with other prognostic factors at the intra-subject level that cannot be explored in a meta-analysis.

Although not consistent across all the M2-like markers, tumor location was also an important factor for the association between CD163 and OS. In contrast to CD68, tumors in the lung and liver showed significantly weaker correlations between CD163 and worse prognosis. Interestingly, CD204 showed a similar trend, with lung and liver carcinomas showing a weaker association with OS. Recent publications studying single cell transcriptomics of the immune microenvironment of lung adenocarcinoma and HCC have revealed prominent macrophage infiltration with more anti-inflammatory (M2 state) features [16, 17]. A subset of HCC-associated macrophages may co-express M1- and M2-associated genes. Lung adenocarcinoma-associated macrophages showed increased CD163 but decreased CD204 gene expression as compared with their healthy lung tissue counterparts. The gene expression

results support what has been described at the protein level in that macrophages are functionally diversified immune cells, widely spread throughout different tumors. The other variable that profoundly moderated CD163 association with OS was the scoring method chosen by the researchers. Macrophage measurements that were taken by hot-spot counting showed stronger HR than those taken by multiple spots where either average or absolute macrophage counting were performed. This indicates the importance of using highly sensitive and objective approaches to measure macrophages. The impact of the scoring strategy has previously been studied for lymphocyte evaluation in triple negative breast cancers with similar results [18]. Bioimage analysis software is a possible solution to standardize IHC quantification. Computer programs, such as QuPath, are an open-source solution for digital pathology and whole slide image analysis and could help to make the quantification of IHC more objective, reliable, and reproducible [19].

Fewer articles studying M1 markers have been published compared with their M2 counterparts. It is worth mentioning that, although non-robustly, the expression of these differentiation markers was associated with better prognosis. At the moment, the scarcity of articles analyzing these markers limits the generalizability of their results.

This meta-analysis has some limitations. First, there was significant data heterogeneity reported by each manuscript. Many manuscripts did not provide relevant information related to the samples studied, the study design, or method used to quantify macrophages. Further efforts should be made in future studies to ensure greater data harmonization. Second, we only included markers measured in tissue by IHC or IF, and therefore our results may not apply to other antibody-based methodologies such as flow cytometry on cell suspensions. The results of this study should be interpreted taking into account that macrophages share some markers with other cell types (IE: dendritic cells) and single-plexed techniques cannot accurately differentiate between these leukocyte populations. As a consequence, the associations may be biased due to the scoring of macrophage-unrelated populations. Third, not all tumor histologies and locations were represented for all the markers analyzed, which limits the statistical power to find associations and the generalizability of results. Finally, a significant amount of residual heterogeneity remained despite including moderators in the models. Although this is probably related to the heterogeneity in the samples and the methods and analyses used across studies that were not possible to take into account in this review, heterogeneity directly correlates with the number of manuscripts included. Therefore, the large number of manuscripts accepted in this review of itself partially explains the high levels of heterogeneity and in no way invalidates the conclusions drawn from this meta-analysis.

With the growing interest in multi-target immunotherapies, there is a real need to refine macrophage measurement. The challenge of inferring the functional properties of macrophages with surrogate surface markers remains unresolved. To date, no single marker can elucidate macrophage polarization. In addition, as stated above, a single marker may be expressed in different macrophage functional states. One possible solution is using multiplexed IF assays that allow the identification of multiple markers in a single cell [20]. Continuous efforts are being made to reliably identify markers able to subclassify macrophages in several clinical contexts including cancer. Recently, macrophage subsets identified by the combination of markers such as HLA-II, CD9, and CD301b have been shown to correlate with functional states in murine models of foreign body reaction [21]. Some of these subsets are also present in non-epithelial tumor models [21]. As yet, these markers have not been studied in human cancer and further research will be needed to assess their role in polarized TAM identification.

The data collected in this review suggest that the available macrophage markers are able to detect subpopulations of macrophages that could potentially influence tumor progression and aggressiveness. It should be noted that treatment with checkpoint inhibitors may alter these associations.

We can conclude that, in line with previous studies, the expression of CD68 in tumors is associated with worse prognosis across all tumors except for colorectal carcinoma [8]. The proposed M2 markers, such as CD163, CD204, and CD206 were associated with worse prognosis. At present, the published data available for other macrophagic markers is too scarce for any firm conclusion to be reached. Future research in the field should take into consideration that certain variables may greatly influence the outcome of the study, namely, the anatomical location of the tumors, the macrophage marker being measured, the compartment (intra-tumor vs. invasive front) where macrophages are scored, and the scoring strategy chosen. Further efforts also are needed to improve and standardize the reporting of information in manuscripts studying macrophage markers.

Acknowledgements We thank Dr Ignacio Ruz Caracuel for his comments and advice regarding article preparation.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Siveen KS, Kuttan G. Role of macrophages in tumour progression. *Immunol Lett.* 2009;123:97–102.
2. Biswas SK, Mantovani A. Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. *Nat Immunol.* 2010;11:889–96.
3. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P. Tumour-associated macrophages as treatment targets in oncology. *Nat Rev Clin Oncol.* 2017;14:399–416.
4. Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, Locati M. The chemokine system in diverse forms of macrophage activation and polarization. *Trends Immunol.* 2004;25:677–86.
5. Shrivastava R, Shukla N. Attributes of alternatively activated (M2) macrophages. *Life Sci.* 2019;224:222–31.
6. Heusinkveld M, van der Burg SH. Identification and manipulation of tumor associated macrophages in human cancers. *J Transl Med.* 2011;9:216.
7. Cassetta L, Pollard JW. Targeting macrophages: therapeutic approaches in cancer. *Nat Rev Drug Disco.* 2018;17:887–904.
8. Zhang Q, Liu L, Gong C, Shi H, Zeng Y, Wang X, et al. Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature. *PLoS ONE.* 2012;7:e50946.
9. Altman DG, Bland JM. How to obtain the confidence interval from a P value. *BMJ.* 2011;343:d2090.
10. Fox J. The R Commander: a basic statistics graphical user interface to R. *J Stat Softw.* 2005;14:1–42.
11. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 2010;36:1–48.
12. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics.* 2000;56:455–63.
13. Gordon M, Lumley T. forestplot: Advanced Forest Plot Using “grid” Graphics. R package version 1.9. 2019.
14. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer, New York; 2016.
15. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339:b2535.
16. Lavin Y, Kobayashi S, Leader A, Amir E-AD, Elefant N, Bigenwald C, et al. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. *Cell.* 2017;169:750–65.
17. Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell.* 2019;179:829–45.
18. McIntire PJ, Irshaid L, Liu Y, Chen Z, Menken F, Nowak E, et al. Hot spot and whole-tumor enumeration of CD8+ tumor-infiltrating lymphocytes utilizing digital image analysis is prognostic in triple-negative breast cancer. *Clin Breast Cancer.* 2018;18:451.
19. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep.* 2017;7:16878.
20. Surace M, DaCosta K, Huntley A, Zhao W, Bagnall C, Brown C, et al. Automated multiplex immunofluorescence panel for immuno-oncology studies on formalin-fixed carcinoma tissue specimens. *J Vis Exp.* 2019;143:e58390.
21. Sommerfeld SD, Cherry C, Schwab RM, Chung L, Maestas DR, Laffont P, et al. Interleukin-36 γ -producing macrophages drive IL-17-mediated fibrosis. *Sci Immunol.* 2019;4:eaax4783.