

An Expert-Annotated Dataset of Bone Marrow Cytology in Hematologic Malignancies (Bone-Marrow-Cytomorphology_MLL_Helmholtz_Fraunhofer)

Summary

Redirection Notice

This page will redirect to https://www.cancerimagingarchive.net/collection/bone-marrow-cytomorphology_mll_helmholtz_fraunhofer/ in about 5 seconds.

The dataset contains a collection of over 170,000 de-identified, expert-annotated cells from the bone marrow smears of 945 patients stained using the May-Grünwald-Giemsa/Pappenheim stain. The diagnosis distribution in the cohort included a variety of hematological diseases reflective of the sample entry of a large laboratory specialized in leukemia diagnostics. Image acquisition was performed using a brightfield microscope with 40x magnification and oil immersion.

Large datasets with a high quality of both data acquisition and annotation are key prerequisites to develop data-driven, computational methods in diagnostic medicine. In the case of bone marrow morphology, a key diagnostic method for a broad range of hematologic diseases, only few datasets are publicly available so far, which are orders of magnitude smaller than the one presented here. Inclusion of our dataset into TCIA provides both medical researchers and bioinformaticians with a public resource for education and algorithm improvement.

All samples were processed in the Munich Leukemia Laboratory (MLL), scanned using equipment developed at Fraunhofer IIS and post-processed using software developed at Helmholtz Munich.

Acknowledgements

- Christian Matek and Carsten Marr acknowledge support from the German National Research foundation (DFG) through grant SFB 1243.
- Carsten Marr has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement number 866411).

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Tissue Slide Images (JPG, 6.8GB)	Download Search (Download and apply the IBM-Aspera-Connect plugin to your browser to retrieve this faspex package)	CC BY 4.0

Click the Versions tab for more info about data releases.

Please contact help@cancerimagingarchive.net with any questions regarding usage.

Detailed Description

Detailed Description

Image Statistics	Pathology Image Statistics
Modalities	Pathology
Number of Patients	945

Number of Images	171,375
Images Size (GB)	6.8

Abbreviations:

ABE	Abnormal eosinophil
ART	Artefact
BAS	Basophil
BLA	Blast
EBO	Erythroblast
EOS	Eosinophil
FGC	Faggott cell
HAC	Hairy cell
KSC	Smudge cell
LYI	Immature lymphocyte
LYT	Lymphocyte
MMZ	Metamyelocyte
MON	Monocyte
MYB	Myelocyte
NGB	Band neutrophil
NGS	Segmented neutrophil
NIF	Not identifiable
OTH	Other cell
PEB	Proerythroblast
PLM	Plasma cell
PMO	Promyelocyte

Citations & Data Usage Policy

Citations & Data Usage Policy

Users must abide by the [TCIA Data Usage Policy and Restrictions](https://wiki.cancerimagingarchive.net). Attribution should include references to the following citations:

Data Citation

Matek, C., Krappe, S., Münzenmayer, C., Haferlach, T., & Marr, C. (2021). **An Expert-Annotated Dataset of Bone Marrow Cytology in Hematologic Malignancies [Data set]**. The Cancer Imaging Archive. <https://doi.org/10.7937/TCIA.AXH3-T579>

Publication Citation

Matek, C., Krappe, S., Münzenmayer, C., Haferlach, T., and Marr, C. (2021). **Highly accurate differentiation of bone marrow cell morphologies using deep neural networks on a large image dataset**. <https://doi.org/10.1182/blood.2020010568>

TCIA Citation

Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. DOI: [10.1007/s10278-013-9622-7](https://doi.org/10.1007/s10278-013-9622-7)

Other Publications Using This Data

TCIA maintains [a list of publications](#) which leverage TCIA data. If you have a manuscript you'd like to add please [contact TCIA's Helpdesk](#).

Versions

Version 1 (Current): Updated 2021/11/12

Data Type	Download all or Query/Filter
Tissue Slide Images (JPG, 6.8GB)	Download Search