

Incidence and Survival of adult Central Nervous System Tumors in the Veneto Region: A Population-Based Registry Study (2016–2020)

Journal of Neuro-Oncology

Alessandra Andreotti, Eliana Ferroni, Stefano Guzzinati, Susanna Baracco, Maddalena Baracco, Emanuela Bovo, Eva Carpin, Antonella Dal Cin, Alessandra Greco, Anna Rita Fiore, Laura Memo, Daniele Monetti, Silvia Rizzato, Jessica Elisabeth Stocco, Carmen Stocco, Sara Zamberlan, Alberto Bosio, Sara Lonardi, Giuseppe Lombardi, Manuel Zorzi.

Corresponding author:

Eliana Ferroni

Epidemiological Department, Azienda Zero – Veneto Region

email: eliana.ferroni@azero.veneto.it

Supplementary Material 1. Procedures for the extraction of the tumor grade from each patient's pathologic report

Pathology reports are lengthy and comprise several sections, frequently with repeated spaces or entries between them. The process begins with a comprehensive pre-processing stage aimed at cleaning and standardizing the input text (Figure 1). This includes the removal of line breaks and excessive whitespace, as well as the identification and isolation of the addendum section, which often contains crucial diagnostic details. Moreover, if multiple grades are present within the same diagnosis, specific criteria have been applied to determine and select the most clinically relevant one. Following text normalization, the system searches for predefined keywords associated with grading terminology. Upon detection of such keywords, a substring of fixed length (34 characters) is extracted, encompassing six characters preceding and twenty-eight characters following the keyword. This targeted extraction ensures that the relevant context surrounding the keyword - typically including the grade - is captured. The final step involves applying a tailored function to the extracted substring to identify and retrieve the grade. This function is designed to recognize various formatting patterns and lexical variants of grade expressions, thereby ensuring accurate and robust extraction across diverse free-text inputs.

This technique was used to extract the tumor grade for both CNS malignancies and cases of meningioma (ICD-O-3 behavior code 0 or 1), incident in the Veneto Region from 2016 to 2020.

Figure 1. Flow chart of the methodology adopted for tumor grade extraction

