Chapter 12
Big Data Analytics and Radiomics to Discover Diagnostics and Therapeutics for Gastric Cancer

Kummetha Jagadish, B. Pratap Naidu, G. Mohana Sheela, Nageswara Rao Reddy Neelapu, and Pallaval Veera Bramhachari

Abstract Cancer is the cause of early death and it is unique. A cancer diagnosis is complicated, and treatment outcomes vary from patient to patient. Improving cancer diagnosis may help in early diagnosis and reduces early deaths. The most common method for the diagnosis of gastrointestinal cancer is gastroscopic imaging. The availability of white light, non-magnifying images, and manual pathological examination are the major drawbacks of the system. Imaging methods like X-Ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Nuclear Medicine (NM) Positron Emission Tomography (PET), and Ultrasound (US) had revolutionized the diagnosis of gastrointestinal cancer. The disadvantage with these radiological images is that they contain more information and content, which is not visible to the clinician's eye. Radiomics is a process of conversion of digital medical images into mineable high-dimensional data. In this chapter, the use of big data in radiomics as a tool for gastrointestinal cancer diagnosis and prognosis is discussed. This provides information and helps in the early detection of gastrointestinal cancer.

Keywords Cancer · Gastrointestinal cancer · Radiological images · Radiomics

K. Jagadish
GVK BIO, Hyderabad, Telangana State, India

B. P. Naidu (✉) · G. M. Sheela · P. Veera Bramhachari
Department of Biotechnology, Krishna University, Machilipatnam, Andhra Pradesh, India

N. R. R. Neelapu
Department of Biochemistry and Bioinformatics, GITAM Institute of Science, Gandhi Institute of Technology and Management (GITAM) (Deemed-to-be-University), Visakhapatnam, Andhra Pradesh, India

P. Veera Bramhachari, N. R. R. Neelapu (eds.), *Recent Advancements in Biomarkers and Early Detection of Gastrointestinal Cancers*, Diagnostics and Therapeutic Advances in GI Malignancies, https://doi.org/10.1007/978-981-15-4431-6_12

12.1 Introduction

Gastric cancer (GC) is the third-highest based on lethality and fourth-highest based on morbidity of all cancers (Rawla and Barsouk 2019). According to WHO GLOBOCAN statistics ~1,033,701 new cases and ~782,000 deaths were recorded for gastric cancer in 2018 (Bray et al. 2018). The most important reason for the situation is the lack of methods for early diagnosis. Early signs of GC are extremely difficult to detect, often bearing a close resemblance to inflammation. When diagnosed, half of the GC patients are in an advanced stage having a 5-year survival rate which is lower than 30% (Rugge et al. 2014). Early detection and proper treatment following precise risk classification are crucial for improving the outcome of gastric cancer. Gastroscopic imaging is a widely used method for the diagnosis of GC. The drawbacks include the availability of white light and non-magnifying images, manual pathological inspection, the inability of the human eye to identify minor lesions from the images, requirement of high-quality, narrow-band imaging (or laser-based), and requirement of magnified images for present image reading algorithms. Recently, computer-aided methods are expected to play an important role in the detection of GC. Development of advanced magnifying endoscopes, deep learning methods, and machine learning methods; and availability of histopathological images enabled reading the weakly labeled images. These advances and developments improved the diagnosis of GC (Ronald 2018). The above methods are used to diagnose GC, but the methods to diagnose GC at very early stages are required.

GC is a disease that evolves due to various genetic and epigenetic alterations. GC originates due to the sequential accumulation of molecular and genetic alterations in stomach epithelial cells. Multidisciplinary diagnostic approaches integrating endoscopy, serology, histology, and molecular profiling are the appropriate approaches for stratification of patients into different GC risk classes. Big data analytics and machine learning methods can bring together the above-mentioned multiple disciplinary diagnostics to help in early diagnosis of GC. The term "big data" refers "to huge amounts of information that can be analyzed by high-performance computers to reveal patterns, trends, and associations." In medical terms, big data includes clinical and genomic data that is derived from patients during diagnostic testing and treatment. Big data analytics can reveal the patterns and relationships among a large amount of data in a single or several data sets. The data analytics uses several techniques like statistics and artificial intelligence to reveal the hidden patterns and rules in big data. Big data analytics is used in a variety of activities or applications, and the application of big data analytics to the gastric cancer diagnosis is an upcoming trend. Recent advances in understanding the molecular mechanisms that mediate GC and big data analytics were promising and paved the path for the development of more effective diagnosis strategies. Extensive research is also carried out in the field of image analysis for diagnosing and identifying GC at the early stages. Recently, Japanese research group successfully used artificial intelligence to diagnose GC (In breakthrough, Japanese researchers use AI to identify early-stage stomach cancer with high accuracy 2018). In this chapter, the recent updates on the role of big data in radiomics, machine learning, or artificial intelligence for early diagnosis of GC are discussed.

12.2 Radiomics

Radiological imaging techniques are powerful noninvasive tools used for detection, differentiation, and diagnosis of different tissue characteristics in patients. Radiologists acquire a huge amount of data by imaging tissues from various views and angles for complete image phenotypes. The imaging methods include X-Ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Nuclear Medicine (NM), Positron Emission Tomography (PET), and Ultrasound (US). Each of these modalities creates tissue contrast based on the differences in the tissue between normal or abnormal. These tissue contrasts are exploited by the radiologist to identify patterns for diagnosis. Radiologists are trained to understand the imaging phenotypes and transcribe these observations to correlate with underlying diseases. Traditionally, these medical images are treated as pictures intended solely for visual interpretation. However, each of the radiological images contains more information content not visible to the clinician's eye. This "hidden" information creates a "radiological texture" which can provide much more information about the tissue of interest than previously thought.

Radiomics is a promising field of medical research that employs a combination of computer-aided deep learning methods and human skills to convert digital medical images into mineable high-dimensional data (Lambin et al. 2012). Then translates the metrics obtained from texture and other features on radiological images. Figure 12.1 represents the basic workflow of radionics with imaging, segmentation, feature extraction, and analysis. Medical or radiological images are generated from various modalities such as X-Ray, CT, MRI, PET, and US. Segmentation is performed to define the tumor region on the radiological images. Then, radiomics employs machine learning methods to extract huge quantities of imaging features like tumor intensity, texture, and shape from radiological images. Radiomics features contain useful spatial and textural information on the grayscale patterns and the correlation between image pixels. These features can be modeled or used for analysis assessed for their prognostic power, or linked with stage, or gene expression (Parekh

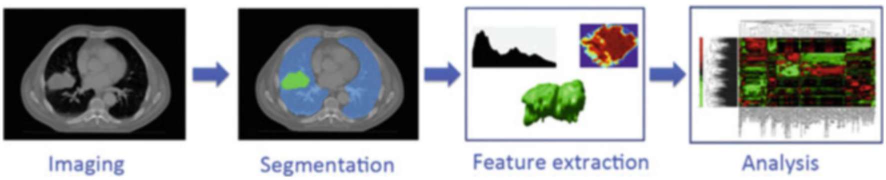

Fig. 12.1 The workflow of radiomics, imaging, segmentation, feature extraction, and analysis of features for prognosis or diagnosis (Parekh and Jacobs 2016)

and Jacobs 2016). Then, this can supplement as an adjunct instrument to discover or predict or decode concealed genetic and molecular traits for decision support, individualized diagnosis, and treatment guidance (Kumar et al. 2012a; Cook et al. 2014; Court et al. 2016; Gillies et al. 2016; Narang et al. 2016; Yip and Aerts 2016; Sala et al. 2017). This information obtained from these radiological images can also be combined with additional OMICS data (genomics, proteomics, metabolomics, and transcriptomics) for further analysis. This branch of study is called radiogenomics, and this is an upcoming technology that is applied for diagnosis and prognosis of multiple cancers (Zinn et al. 2011).

12.3 Big Data in Radiomics for Diagnosis and Prognosis of Gastric Cancer

Radiomics is an upcoming technology that can be used in general to predict gastric cancer and plan for a course of treatment. Recently, multiple studies have discussed the possibility of using radiomics and artificial intelligence for diagnosis and prognosis of gastric cancer (Jiang et al. 2018a, b; Li et al. 2018a, b; Keek et al. 2018; Acharya et al. 2018).

12.3.1 Radiomics in Preoperative Prediction of Lymph Node Metastasis

Feng et al. (2019) developed and validated an automatic decision support system (DSS) for preoperative reporting of the risk for lymph node metastasis in GC. The clinical and imaging data were analyzed using a machine learning-based approach. The clinical, pathological, and CT imaging data of 490 patients diagnosed with GC was collected. Standard gastric contrast-enhanced CT scans of the same patients were also obtained within 10 days of surgery and all gastric CT studies were performed using a 64-slice scanner. Of the 490 patients, 297 were reported with LN metastasis and also with the metastatic rate of 60.6%. Thirteen relevant radiomics features were selected, ranked, and modeled using a support vector machine (SVM) classifier based on 326 training and validation data sets. A model test was performed independently with a test set size (n) 164. The comparison was made between the Clinical Decision Support System (CDSS) and the conventional staging criterion performed by two expert radiologists for the diagnostic performance of CDSS. The DSS was better able to predict LN metastasis (accuracy 76.4%) than the conventional staging (accuracy 71.3%). Automatic DSS employing SVM classifier was able to predict LN status in patients with GC based on 13 radiomics features.

12.3.2 *Radiomics to Predict Prognosis and Benefit from Chemotherapy*

Jiang et al. (2018b) developed and validated a radiomics signature for the prediction of gastric cancer and the benefits of chemotherapy. A sample of 1591 patients histologically confirmed with gastric adenocarcinoma, and standard unenhanced and contrast-enhanced abdominal CT performed within 30 days was included for analysis. Radiomics signature and radiomics nomograms were generated from the sample. The lasso-cox regression model was performed on 228 patients to generate radiomics signature based on 19 selected features. Radiomics nomograms integrated with radiomics signature were constructed on TNM staging. Radiomics signature was able to predict patients with stage II and III of GC and may benefit chemotherapy.

Both the case studies could predict gastric cancer with very high accuracy. Though good progress is seen with radiomics, some challenges also exist. The major challenges to be addressed for analyzing the data are standard image acquisition methods, image reconstruction methods, optimized algorithmic approaches, and statistical approaches. Databases like the National Biomedical Imaging Archive (NBIA) (Nicholas et al. 2012), Cancer Imaging Program (CIP) (Dobranowski et al. 2014), and The Cancer Imaging Archive (TCIA) (Clark et al. 2013) are available. Developing integrated radiomics images with defined rules might help in addressing the challenges with image acquisition (Kumar et al. 2012b). However, these computer- assisted clinical decision-making methods require further external, multicenter, and evidence-based validation. Further, these applications may serve as a tool for personalized diagnosis and guidance for treatment (Parekh and Jacobs 2016).

12.4 Conclusion and Future Perspectives

In conclusion, radiomics analysis uses a machine-learning approach to provide an alternative to conventional radiologic methods. This, in turn, changes the facet of clinical decision-making of the present and future generations of patients suffering from gastric cancer. The future direction of radiomics includes correlating and integrating OMICS data with radiomics features extracted from radiological images and integrate them to create a more efficient and robust prognostic model. This, well aid clinicians in regular practice, personalized medicine, and paves new direction for the cancer diagnosis and prognosis (Mazurowski 2015; Rutman and Kuo 2009).

Acknowledgments KVJ is grateful to GVK BIO, Hyderabad, for providing necessary facilities to carry out the research work and for extending constant support. NNRR is grateful to GITAM (Deemed-to-be-University) for providing necessary facilities to carry out the research work and for extending constant support. Prathap Naidu, Dr. Mohana Sheela, and Dr. PVBC are grateful to Krishna University for providing the necessary facilities to carry out the research work and for extending constant support.

Conflict of Interest Statement The authors declare that there is no potential conflict of interest.

References

Acharya UR, Hagiwara Y, Sudarshan VK, Chan WY, Ng KH (2018) Towards precision medicine: from quantitative imaging to radiomics. J Zhejiang Univ Sci B 19(1):6–24. https://doi.org/10.1631/jzus.B1700260

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394–424

Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26(6):1045–1057. https://www.cancerimagingarchive.net/

Cook GJR, Siddique M, Taylor BP, Yip C, Chicklore S, Goh V (2014) Radiomics in PET: principles and applications. Clin Transl Imaging 2(3):269–276

Court LE, Fave X, Mackin D, Lee J, Yang J, Zhang L (2016) Computational resources for radiomics. Transl Cancer Res 5(4):340–348

Dobranowski J, Melamed S, Langer D, Bedford C (2014) The cancer imaging program quality framework at cancer care Ontario: the first five years. J Clin Oncol 32(30_suppl):244. https://imaging.cancer.gov/

Feng QX, Liu C, Qi L, Sun SW, Song Y, Yang G, Zhang YD, Liu XS (2019) An intelligent clinical decision support system for preoperative prediction of lymph node metastasis in gastric cancer. J Am Coll Radiol 16(7):952–960

Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278(2):563–577

In breakthrough, Japanese researchers use AI to identify early-stage stomach cancer with high accuracy (2018). https://www.japantimes.co.jp/news/2018/07/22/national/science-health/japanese-researchers-use-ai-identify-early-stage-stomach-cancer-high-accuracy/#.XFqIPFwza70. Accessed 12 Oct 2019

Jiang Y, Chen C, Xie J, Wang W, Zha X, Lv W, Chen H, Hu Y, Li T, Yu J, Zhou Z, Xu Y, Li G (2018a) Radiomics signature of computed tomography imaging for prediction of survival and chemotherapeutic benefits in gastric cancer. EBioMedicine 36:171–182

Jiang Y, Yuan Q, Lv W, Xi S, Huang W, Sun Z, Chen H, Zhao L, Liu W, Hu Y, Lu L, Ma J, Li T, Yu J, Wang Q, Li G (2018b) Radiomic signature of 18F fluorodeoxyglucose PET/CT for prediction of gastric cancer survival and chemotherapeutic benefits. Theranostics 8(21):5915–5928

Keek SA, Leijenaar RT, Jochems A, Woodruff HC (2018) A review on radiomics and the future of theranostics for patient selection in precision medicine. Br J Radiol 91(1091):20170926. https://doi.org/10.1259/bjr.20170926

Kumar V, Gu Y, Basu S, Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ (2012a) Radiomics: the process and the challenges. Magn Reson Imaging 30(9):1234–1248

Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ (2012b) Radiomics: the process and the challenges. Magn Reson Imaging 30(9):1234–1248. https://doi.org/10.1016/j.mri.2012.06.010

Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48(4):441–446

Li Z, Zhang D, Dai Y, Dong J, Wu L, Li Y, Cheng Z, Ding Y, Liu Z (2018a) Computed tomography-based radiomics for prediction of neoadjuvant chemotherapy outcomes in locally advanced gastric cancer: a pilot study. Chin J Cancer Res 30(4):406–414

Li W, Zhang L, Tian C, Song H, Fang M, Hu C, Zang Y, Cao Y, Dai S, Wang F, Dong D, Wang R, Tian J (2018b) Prognostic value of computed tomography radiomics features in patients with gastric cancer following curative resection. J Eur Radiol 29(6):3079–3089

Mazurowski MA (2015) Radiogenomics: what it is and why it is important. J Am Coll Radiol 12 (8):862–866

Narang S, Lehrer M, Yang D, Lee J, Rao A (2016) Radiomics in glioblastoma: current status, challenges and potential opportunities. Transl Cancer Res 5(4):383–397

Nicholas A, Mulhern P, Siegel E (2012) The National Biomedical Imaging Archive: a repository of advanced imaging information. J Nucl Med 53(Suppl 1):1009. https://imaging.nci.nih.gov/ncia/login.jsf

Parekh V, Jacobs MA (2016) Radiomics: a new application from established techniques. Expert Rev Precis Med Drug Dev 1(2):207–226

Rawla P, Barsouk A (2019) Epidemiology of gastric cancer: global trends, risk factors and prevention. Prz Gastroenterol 14(1):26–38

Ronald S (2018) The impact of deep learning and artificial intelligence on radiology. In: 1st conference on medical imaging with deep learning (MIDL 2018), Amsterdam, The Netherlands, pp 6

Rugge M, Capelle LG, Fassan M (2014) Individual risk stratification of gastric cancer: evolving concepts and their impact on clinical practice. Best Pract Res Clin Gastroenterol 28 (6):1043–1053

Rutman AM, Kuo MD (2009) Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. Eur J Radiol 70(2):232–241

Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, Weigelt B, Vargas HA (2017) Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. Clin Radiol 72(1):3–10

Yip SSF, Aerts HJWL (2016) Applications and limitations of radiomics. Phys Med Biol 61(13): R150–R166

Zinn PO, Mahajan B, Sathyan P, Singh SK, Majumder S, Jolesz FA, Colen RR (2011) Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. PLoS One 6(10):e25451