Predicting Cervical Cancer Outcomes: Statistics, Images, and Machine Learning

Wei Luo*

Department of Radiation Medicine, University of Kentucky, Lexington, KY, United States

Cervical cancer is a very common and severe disease in women worldwide. Accurate prediction of its clinical outcomes will help adjust or optimize the treatment of cervical cancer and benefit the patients. Statistical models, various types of medical images, and machine learning have been used for outcome prediction and obtained promising results. Compared to conventional statistical models, machine learning has demonstrated advantages in dealing with the complexity in large-scale data and discovering prognostic factors. It has great potential in clinical application and improving cervical cancer management. However, the limitations of prediction studies and prediction models including simplification, insufficient data, overfitting and lack of interpretability, indicate that more work is needed to make clinical outcome prediction more accurate, more reliable, and more practical for clinical use.

Keywords: cervical cancer, clinical outcome prediction, statistical model, machine learning, medical image, radiomics

INTRODUCTION

Cancer treatment is one of the most complicated and challenging tasks in medicine. Although cancer survival rate has been significantly improved for the last decades with the introduction of new drugs, technologies and techniques, there are still uncertainties on the effect of those advances on clinical outcomes. The information of clinical outcomes is critical for the evaluation of treatment effectiveness and optimization of treatment strategies. Clinical outcomes usually are not available until enough clinical data have been accumulated following up a large number of patients for long periods. To know clinical outcomes more quickly so that treatment can be improved or adjusted timely, accurate prediction of clinical outcomes is expected. There are two approaches used for clinical outcome prediction. One is to use radiobiological models including tumor control probability (TCP) model, normal tissue complication probability (NTCP) model, and equivalent uniform dose (EUD). The other is to build statistical models utilizing all the information that is relevant to disease prognosis such as demographics, laboratory tests, images, and dosimetry, to find the relationship between those factors and clinical outcomes. The more data is used, the more accurate the prediction would be. In this regard, artificial intelligence especially machine learning (ML) has a great capacity to process huge and complex data and thus has been used in many areas including medicine. Recently, ML has been introduced into radiation oncology to predict clinical outcomes (Kang et al., 2015; Luo et al., 2020).

Cervical cancer is the third most common cancer and a leading cause to death for women worldwide (Ferlay et al., 2019; Rebecca, 2020). It is one of a few cancers that were first treated with radiation therapy successfully (Mazeron and Gerbaulet, 1998). The treatment of cervical cancer is also one of the most complex and challenging cancer management tasks and may involve all three
cancer treatment modalities (surgery, chemotherapy, and radiation therapy (RT)) and all radiation therapy techniques (external beam radiation therapy (EBRT), intracavitary/interstitial brachytherapy (BT), high dose rate (HDR)/low dose rate (LDR) brachytherapy, and permanent seed implant). This paper does not intend to provide a comprehensive review of cervical cancer outcome predictions, but mainly focuses on the prediction results with different methods, the efficacy and limitations of prediction associated with radiation therapy.

Reported Clinical Outcomes

Actual clinical outcomes are directly derived from the results obtained following up patients. Numerous studies have revealed cervical cancer survival rates for different International Federation of Gynecology and Obstetrics (FIGO) stages (I–IV) and different treatment techniques. In the United States, the 5-year survival rates of cervical cancer patients ranged from 17 to 92% with the all-stage rate of 66% according to the American Cancer Society (American Cancer Society, 2020). Surgery, chemotherapy, and radiation therapy are the treatment options for cervical cancer. The clinical outcomes are associated with treatment modalities and FIGO stages. The actual 5-year survival rates have been reported and are summarized in Table 1 (Brunschwig, 1968; Kim et al., 1988; Landoni et al., 1997; Joslin et al., 2001; Eifel et al., 2004). Severe complications were also reported for 9% of patients with radiation therapy alone (Podczaski et al., 1990) and 20% of patients with chemoradiotherapy (Small et al., 2011). Those reported results were summaries of previous clinical data, but not predictions of clinical outcomes. Mathematical models can establish quantitative relationship between disease-related factors and outcomes and thus predict clinical outcomes based on identified prognostic factors or predictors.

Outcome Prediction Using Image Analysis

Radiation therapy heavily relies on medical imaging. Various three-dimensional (3D) imaging techniques such as computerized tomography (CT), nuclear magnetic resonance imaging (MRI) and positron emission tomography (PET) have been widely used for cervical cancer diagnosis and treatment. Those images may also contain the information about clinical outcomes. By analyzing the F-18 fluorodeoxyglucose (FDG) pretreatment images of 248 cervical cancer patients staged from IA2 to IVB and using CPHR, a study reported that the maximal standardized uptake value (SUVmax) that quantifies cervical tumor uptake of FDG is associated with treatment response and prognosis in cervical cancer patients and gave better outcome prediction than lymph node status, stage, or tumor volume (Kidd et al., 2007). The results showed that the overall survival rate at 5 years was 95% for patients with an SUVmax ≤5.2, 70% for patients with an SUVmax from >5.2 to ≤13.3, and 44% for patients with an SUVmax >13.3.

Recently, radiomics has been introduced as a powerful tool to extract huge and complex image features from PET/CT and MRI images for prediction of cervical cancer clinical outcome. It was reported that radiomics features could contribute to prognoses in cervical cancer (Lucia et al., 2018). Using CPHR, two textural features, Grey Level Non Uniformity gray-level run-length matrix (GLRLM) in PET and Entropy gray-level co-occurrence matrix GLCM in ADC maps from DWI MRI, were identified as independent prognostic factors. They were significantly stronger correlated with prognoses than clinical parameters, with an accuracy of 94% for predicting recurrence and 100% for predicting lack of loco-regional control compared with ~50–60% accuracy with clinical parameters. It was also found that the high gray-level run emphasis (HGRE) derived from GLRLM and used to measure high SUV distribution can serve as a predictor (Chen et al., 2018a). This study included 142 cervical cancer patients who had took 18F-FDG PET/CT for pretreatment staging and treated with EBT and intracavitary brachytherapy as well as concurrent chemotherapy. The binary logistic regression model was used to identify the independent prognostic factors among all the radiomic features and predict clinical outcomes. The log-rank test and CPHR analysis were performed to examine the effects of explanatory variables on outcome endpoints including overall survival, progression-free
TABLE 2 | The results from machine learning.

| Algorithm                  | Study                  | Accuracy | Sensitivity | Specificity | AUC  | End point         |
|----------------------------|------------------------|----------|-------------|-------------|------|------------------|
| PNN                        | Obrzut et al. (2017)   | 0.9      | 1.0         |             |      | 5 year-survival  |
|                            | Obrzut et al. (2019)   | 0.9      | 0.7         |             |      | 10-year survival |
| Network in network         | Shen et al. (2019)     | 89.0     | 71.0        | 93.0        | 0.70 | Metastasis       |
|                            |                       | 87.0     | 77.0        | 90.0        |      | Pelvic toxicity  |
| CNN                       | Zhen et al. (2017)     | 72.0     | 59.0        |             | 0.70 | Rectal toxicity  |
| SVM                       | Chen et al. (2018b)    | 87.8     | 79.9        |             | 0.91 | Rectal toxicity  |
| SVM                       | Tian et al. (2019)     | 97.1     | 88.5        |             | 0.90 | Radiation-induced fistula |

Outcome Prediction Using Machine Learning

To the author’s knowledge, ML was first used to predict overall survival for 134 cervical cancer patients in 2002, using an artificial neural network model (ANN) including 11 prognostic factors (age, performance status, hemoglobin, total protein in serum, FIGO stage, histological type, histological grading at 30 Gy, histological grading at 40 Gy, histological grading at the end of therapy, cytological grading at 30 Gy, cytological grading at 40 Gy, cytological grading at the end of therapy) (Ochi et al., 2002). The predicted survival result was able to achieve an area under the receiver operating characteristic (ROC) curve (AUC) of 0.7782. A more recent study included 102 patients with cervical cancer staged as IA2-IIB, selected 23 demographic and tumor-related parameters, and collected perioperative data of each patient (Obrzut et al., 2017). The study predicted the 5-year survival rate using six machine learning methods: the probabilistic neural network (PNN), multilayer perceptron network (MLP), gene expression programming classifier (GEP), support vector machines algorithm (SVM), radial basis function neural network (RBFNN) and k-Means algorithm. Compared with other models, PNN provided the best prediction with an accuracy of 0.892 and sensitivity of 0.975. PNN was further used to predict the 10-year survival for the same cohort and also achieved high predictability (Obrzut et al., 2019).

Deep-learning (DL) has also been introduced for outcome prediction. A neural network model was implemented to predict survival utilizing clinicolaboratory variables among recurrent cervical cancer patients (Matsuo et al., 2017). The study tried to find among 13 clinicolaboratory variables the predictors for life expectancy in 157 recurrent cervical cancer patients. Those variables included age, body habitus change, pain score, blood pressure, and heart rate, white blood cell, hemoglobin, platelet, bicarbonate, blood urea nitrogen, creatinine, and albumin. The results showed that the 3-month survival decrease was associated with older age, decreasing albumin level, decreasing body mass index, increasing pain score, decreasing systolic blood pressure, decreasing white blood cell count, increasing platelet counts, and decreasing hemoglobin levels. This study group further predicted survival rate for 768 cervical cancer patients using the same DL model with 40 features that included patient demographics, vital signs, laboratory test results, tumor characteristics, and treatment types (Matsuo et al., 2019). They showed that the results of DL were better than that of CPHR.

In a recent study, a DL model called network in network was developed to predict treatment failures including local relapse and distant metastasis based on the analysis of the PET/CT images (Shen et al., 2019). The prediction of local relapse and distant metastasis obtained reasonable accuracy, sensitivity, and specificity. (Table 2) Four groups of radiomic features were also calculated, but none of the radiomic features was able to predict distant metastasis in this study.

ML is also able to predict treatment complications. A retrospective study applied the convolutional neural network (CNN) algorithm to analyze rectum dose distribution and predict rectum complications (Zhen et al., 2017). The study included 42 cervical cancer patients treated with EBRT combined with BT. The results showed that the texture features derived from the rectum surface dose map can generate better predictive performance than the volume parameters D0.1/1/2cc that are prescribed for dose constrains, in terms of sensitivity, specificity and AUC. The same research group applied the SVM algorithm to predict rectal toxicity for the same patient cohort and also achieved higher sensitivity, specificity, and AUC when compared with D0.1/1/2cc (Table 2) (Chen et al., 2018b).

The radiation-induced fistula is a concern for treating advanced gynecological (GYN) malignancies using radiation therapy. Another SVM model was developed to predict the risk of fistula formation caused by radiation therapy (Tian et al., 2019). The study included 35 gynecological cancer patients treated with interstitial BT. The model used the features of mixed data types that might be correlated to fistula formation, and included patient demographics, patient health status, tumor characteristics, additional invasive procedures, and dosimetric parameters. The predicted outcomes achieved a high prediction accuracy as shown in Table 2.

DISCUSSION

Accurate prediction of clinical outcomes would guide treatment to focus on specific prognostic factors and optimize the treatment scheme for each patient. The prediction of cervical cancer...
outcomes is one of the most challenging tasks as the management of cervical cancer involves the most complicated cancer treatment strategies. The studies reviewed in this paper have utilized models to discover many new prognostic factors such as tumor diameter, histological subtype, FDG SUVmax, radiometric features, and clinicolaboratory variables, and establish the relationships between those factors and clinical outcomes. Therefore, clinical outcomes can be accurately predicted. But the accuracy of prediction is related to models and algorithms. Several models performed very well in the studies. For example, CPHR predicted the 5-year survival rates 80% (I), 68% (IIA), and 59% (IIB) (Eifel et al., 2009), which were comparable to the reported results of 83.2% (I) and 68.9% (II) (Kim et al., 1988). Also, several DL models gave high accuracy predictions (Table 2). Such promising results have indicated that model-based outcome prediction has great potential for clinical applications.

The models used for prediction can be categorized into conventional statistical models and ML models. Conventional statistical models include the linear regression, the logistic regression, and CPHR. CPHR is one of most commonly used models for outcome prediction. It models relative hazards treating all the relevant factors proportionally. It can determine which factor is the most influential. But the proposed proportionality or linearity may not be valid because many prognostic factors are not linear and interact with each other. Thus the performance of prediction may not be ideal. In contrast, ML is able to deal with complex and non-linear relations in the data. Especially, it is able to learn feature representations automatically from raw data without direct feature engineering. Overall, ML outperformed statistical models in cervical cancer outcome prediction (Matsuo et al., 2017; Luo et al., 2019; Matsuo et al., 2019; Tian et al., 2019).

However, there was also evidence that the superiority of ML in outcome prediction is not always supported (Christodoulou et al., 2019). In addition, the ML models and algorithms have their own limitations, notably, overfitting (Zhen et al., 2017), and lack of interpretability (Luo et al., 2019; Luo et al., 2020). Overfitting would undermine predictive performance. Lack of interpretability would hinder the use of ML. ML works like a "black box" due to the complex algorithms. It is not easy to understand how it works and the predicted outcomes are not easy to understand as well. For instance, some predictors such as Albumin level were identified as significant prognostic factors by CPHR, but not by the DL (Matsuo et al., 2017). Thus, the prediction using ML may not be as convincing or well accepted as that using conventional models that are explicitly formulated. Furthermore, it is difficult to catch bugs or errors if they occur. Development of independent validation methods may help resolve this issue.

It should also be realized that the studies reviewed in this paper have limitations as well. First of all, most prediction studies did not have enough data, which would reduce the accuracy of the predictions. Secondly, most studies did not distinguish between treatment modalities and techniques. The treatment of cervical cancer involves almost all available cancer treatment modalities and techniques. Each modality and technique play specific roles and has different contributions to clinical outcomes. For example, LDR brachytherapy led to the 4-year disease-free survivals of 87, 66, and 28% for FIGO stages I, II, and III, respectively, (Coia et al., 1990), while HDR was able to achieve the 5-year survival of 94.4, 62, and 37.2%, for state I, II, III, respectively (Utley et al., 1984). Thus, the impact of different techniques on the outcomes should be determined separately and weighted in the prediction models. More attention should be paid to brachytherapy as brachytherapy is a major and complex treatment modality for cervical cancer. Especially, brachytherapy is sensitive to radiobiological effect. Radiobiological effect such as, dose-rate effect should be included in prediction models. Finally, most studies were limited to a single institution and small number of patients, and the results may have bias and significant uncertainties. The predicted outcomes are expected to be comparable to the actual outcomes independently derived from clinical trials or actual patient records.

CONCLUSION

The prediction of cervical cancer outcomes utilizing statistical models, images, and ML has produced promising results. Particularly, ML has capacity to process complex and non-linear relations in large-scale data, discover new prognostic factors, and perform predictions. It has great potential in clinical applications. However, more work is needed to make ML practical and reliable for clinical use. Future studies may include development of new methods and algorithms to minimize the effect of data scarcity, differentiating treatment modalities and techniques in prediction and evaluating individual contributions to clinical outcomes, and independent validation of machine learning algorithms.

AUTHOR CONTRIBUTIONS

The author is fully responsible for the design of the study and writing of the paper.

REFERENCES

American Cancer Society (2020). Cancer Facts and Figures. New York, NY: Oxford University Press.

Brunschwig, A. (1968). The Surgical Treatment of Cancer of the Cervix: Stage I and II. Am. J. Roentgenology 102, 147–151. doi:10.2214/ajr.102.1.147

Chen, J., Chen, H., Zhong, Z., et al. (2018). Investigating Rectal Toxicity Associated Dosimetric Features with Deformable Accumulated Rectal Surface Dose Maps for Cervical Cancer Radiotherapy. Radiat. Oncol. 13, 125. doi:10.1186/s13014-018-1068-0

Chen, S.-W., Shen, W.-C., Hsieh, T.-C., Liang, J.-A., Hung, Y.-C., Yeh, L.-S., et al. (2018). Textural Features of Cervical Cancers on FDG-PET/CT Associate with Survival and Local Relapse in Patients Treated with Definitive Chemoradiotherapy. Sci. Rep. 8, 11859. doi:10.1038/s41598-018-30336-6

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., and Van Calster, B. (2019). A Systematic Review Shows No Performance Benefit of
Machine Learning over Logistic Regression for Clinical Prediction Models. J. Clin. Epidemiol. 110, 12–22. doi:10.1016/j.jclinepi.2019.02.004

Coia, L., Won, M., Lanciano, R., Marcial, V. A., Martz, K., and Hanks, G. (1990). The Patterns of Care Outcome Study for Cancer of the Uterine Cervix, Results of the Second National Practice Survey. Cancer 66, 2451–2456. doi:10.1002/1097-0142(19901215)66:12<2451::aid-cncr2820661202>3.0.co;2-5

Eifel, P. J., Jhinjuran, A., Levenback, C. F., and Tucker, S. (2009). Predictive Value of a Proposed Subclassification of Stages I and II Cervical Cancer Based on Clinical Tumor Diameter. Int. J. Gynecol. Cancer 19, 2–7. doi:10.1111/j.1078-6346.2008.01071.x

Eifel, P. J., Winter, K., Morris, M., Levenback, C., Grigsby, P. W., Cooper, J., et al. (2004). Pelvic Irradiation with Concurrent Chemotherapy versus Pelvic and Para-Aortic Irradiation for High-Risk Cervical Cancer: An Update of Radiation Therapy Oncology Group Cervix Protocol. J. Clin. Oncol. 22, 872–880. doi:10.1200/jco.2004.07.197

Eoh, K. J., Kim, H. J., Lee, J.-Y., Nam, E. J., Kim, S., Kim, S. W., et al. (2017). Upregulation of Homeobox Gene Is Correlated with Poor Survival Outcomes in Cervical Cancer. Oncotarget 8, 84396–84402. doi:10.18632/oncotarget.21041

Felway, J., Colombet, M., Soejmomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., et al. (2019). Estimating the Global Cancer Incidence and Mortality in 2018: GLOBOCAN Sources and Methods. Int. J. Cancer 144, 1941–1953. doi:10.1002/ijc.31937

Joslin, C. (2001). "High Dose Rate Brachytherapy for Treating Cervical Cancer," in Principles and Practice of Brachytherapy Using after Loading Systems. Editors C. Joslin, A. Flynn, and E. Hall (New York: Arnold).

Kang, J., Schwartz, R., Flickinger, J., et al. (2015). Machine Learning Approaches for Predicting Radiation Therapy Outcomes: a Clinician’s Perspective. Int. J. Radiat. Oncol. Biol Phys 93 (No. 5), 1127–1135. doi:10.1016/j.ijrobp.2015.07.2286

Kidd, E. A., Siegel, B. A., Dehdashti, F., and Grigsby, P. W. (2007). The Patterns of Care Outcome Study for Cancer of the Uterine Cervix, Results of the Second National Practice Survey. Cancer 66, 2451–2456. doi:10.1002/1097-0142(19901215)66:12<2451::aid-cncr2820661202>3.0.co;2-5

Luo, Y., Chen, S., and Valdes, G. (2020). Machine Learning over Logistic Regression for Clinical Prediction Models. J. Clin. Epidemiol. 110, 12–22. doi:10.1016/j.jclinepi.2019.02.004

Luo, Y., Won, M., Lanciano, R., Marcial, V. A., Martz, K., and Hanks, G. (1990). The Patterns of Care Outcome Study for Cancer of the Uterine Cervix, Results of the Second National Practice Survey. Cancer 66, 2451–2456. doi:10.1002/1097-0142(19901215)66:12<2451::aid-cncr2820661202>3.0.co;2-5

Matsu, K., Purushotham, S., Moeini, A., Li, G., Machida, H., Liu, Y., et al. (2017). A Pilot Study in Using Deep Learning to Predict Limited Life Expectancy in Women with Recurrent Cervical Cancer. Am. J. Obstet. Gynecol. 217, 703–705. doi:10.1016/j.ajog.2017.08.012

Mazeron, J. J., and Gerbaulet, A. (1998). The Centenary of Discovery of Radium. Radiother. Oncol. 49, 205–216. doi:10.1016/s0167-8140(98)00143-1

Obrutz, B., Kusy, M., Semczuk, A., et al. (2017). Prediction of 5-year Overall Survival in Cervical Cancer Patients Treated with Radical Hysterectomy Using Computational Intelligence Methods. BMC Cancer 17, 840. doi:10.1186/s12885-017-3806-3

Obrutz, B., Kusy, M., Semczuk, A., Obrutz, M., and Klaska, J. (2019). Prediction of 10-year Overall Survival in Patients with Operable Cervical Cancer Using a Probabilistic Neural Network. J. Cancer 10 (18), 4189–4195. doi:10.7150/jca.39345

Ochi, T., Murase, K., Fujii, T., Kawamura, M., and Ikezoe, J. (2002). Survival in Cervical Cancer Patients Treated by Radiation Therapy Alone. J. Clin. Oncol. 7, 0294–0300. doi:10.1016/s1049-7323(02)00043

Podczaski, E., Styrkar, J. A., Kaminski, P., Niedziba, B., Larson, J., Degest, K., et al. (1990). Extended-field Irradiation for Carcinoma of the Cervix. Cancer 66, 251–258. doi:10.1002/1097-0142(19900715)66:2<251::aid-cncr2820661202>3.0.co;2-e

Rebecca, L. (2020). Siegel, Cancer Statistics. Ca Cancer J. Clin. 70, 1. doi:10.3322/cac.21590

Shen, W.-C., Chen, S.-W., Wu, K.-C., Hsieh, T.-C., Liang, J.-A., Hung, Y.-C., et al. (2019). Prediction of Local Relapse and Distant Metastasis in Patients with Definitive Chemoradiotherapy-Treated Cervical Cancer by Deep Learning from [18F]-Fluorodeoxyglucose Positron Emission Tomography/computed Tomography. Eur. Radiol. 29, 6741–6749. doi:10.1007/s00330-019-06265-x

Small, W., Jr, Winter, K., Levenback, C., Iyer, R., Hymes, S. R., Jhingran, A., et al. (2011). Extended-field Irradiation and Intraoperative Brachytherapy Combined with Cisplatin and Amifostine for Cervical Cancer with Positive Para-Aortic or High Common Iliac Lymph Nodes: Results of Arm II of Radiation Therapy Oncology Group (RTOG) 0116. Int. J. Gynecol. Cancer 21, 1266–1275. doi:10.1097/IGC.0b013e3282c7269

Tian, Z., Yen, A., Zhou, Z., Shen, C., Albuquerque, K., and Hrycushko, R. (2019). A Machine-Learning-Based Prediction Model of Fistula Formation after Interstitial Brachytherapy for Locally Advanced Gynecological Malignancies. Brachytherapy 18, 530–538. doi:10.1016/j.bray.2019.04.004

Ulley, J. F., von Essen, C. F., Horn, R. A., and Moeller, J. H. (1984). High-dose-rate Afterloading Brachytherapy in Carcinoma of the Uterine Cervix. Int. J. Radiat. Oncol.*Biology*Physics 10, 2259–2263. doi:10.1016/0360-3016(84)90231-1

Zhen, X., Chen, J., Zhong, Z., Hrycushko, B., Zhou, L., Jiang, S., et al. (2017). Deep Learning in Radiation Therapy: A Review.J. Radiat. Oncol Biol Phys. Med. 62, 8246–8263. doi:10.1088/1361-6556/aa8d09

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Copyright © 2021 Luo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.