Research article

A data assimilation framework to predict the response of glioma cells to radiation

Junyan Liu[1], David A. Hormuth II[2,3], Jianchen Yang[1] and Thomas E. Yankeelov[1,2,3,4,5,6,*]

[1] Department of Biomedical Engineering, The University of Texas at Austin, 107 W Dean Keeton Street Stop C0800, Austin, TX 78712, USA

[2] Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, 201 E. 24th Street POB 4.102 Stop C0200, Austin, TX 78712, USA

[3] Livestrong Cancer Institutes, The University of Texas at Austin, 1601 Trinity St. Bldg. B Mail Stop Z1100, TX 78712, USA

[4] Department of Diagnostic Medicine, The University of Texas at Austin, 1601 Trinity St Bldg. B Stop Z0300, Austin, TX 78712, USA

[5] Department of Oncology, The University of Texas at Austin, 1601 Trinity St Bldg. B, Austin, TX 78712, USA

[6] Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, PO Box 301402, Houston, TX, 77230-1402, USA

* **Correspondence:** Email: thomas.yankeelov@utexas.edu; Tel: +15124713604; Fax: +15124718694.
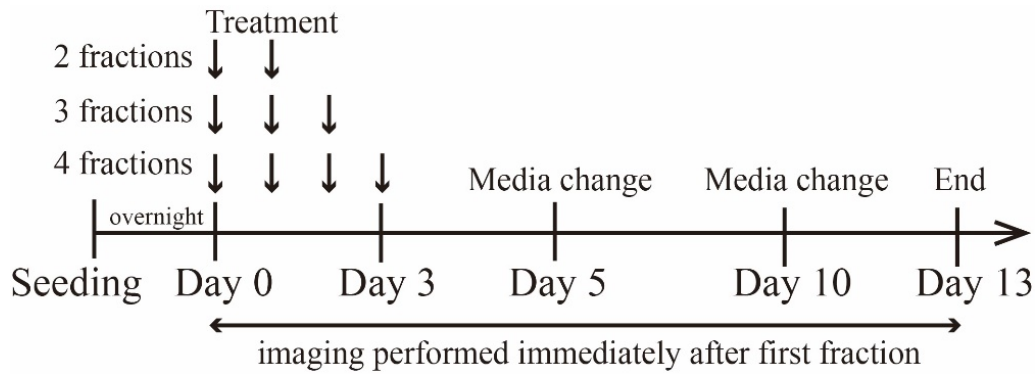
Supplementary

Figure S1. Radiation treatment schedule. Cells are seeded and incubated overnight before treatment. On Day 0, the cells are irradiated with either a total of 16 Gy or 20 Gy *via* different schedules. In the 16 Gy group, the cells receive either two fractions of 8 Gy, three fractions of 5.3 Gy, or four fractions of 4 Gy. In the 20 Gy group, the cells receive either two fractions of 10 Gy, three fractions of 2 Gy, or four fractions of 5 Gy. All fractions are separated by 24 hours. We refresh the cell culture media on Days 5 and 10. Microscopy images are collected continuously right after the first fraction on Day 0 until the end of Day 13.
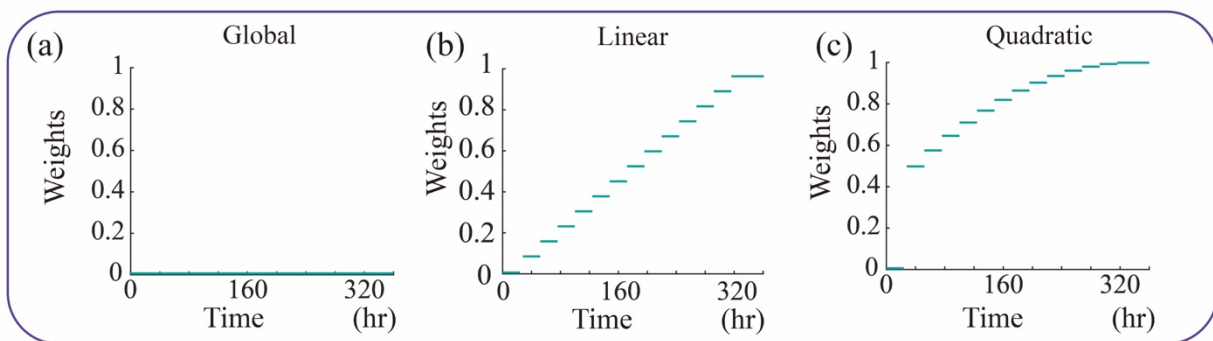

Figure S2. Illustrating different methods of weighting the individual parameters. In panel (a), global prediction is performed by using only Xpop, which is equal to setting the weights on Xind,i to 0 throughout the time course. In panel (b), we increase the weights linearly as in Eq (11). In panel (c), we increase the weights quadratically via Eq (13). Compared to the linear weighting, the quadratic weighting scheme emphasizes the sample-specific measurements earlier in the time course.
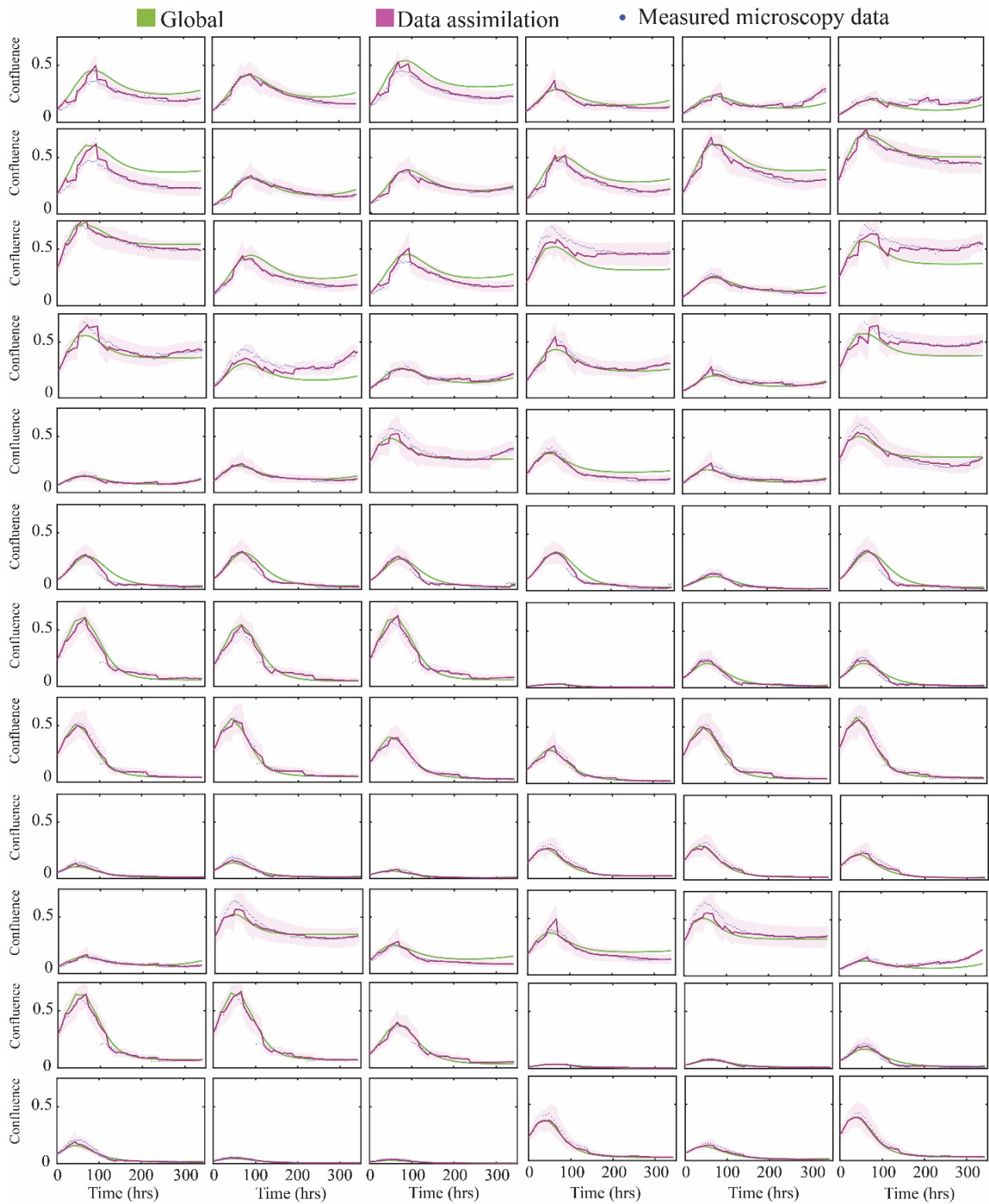
Figure S3. The data assimilation prediction versus global prediction for all 9L samples from the validation group. See caption to Figure 6 for an explanation of the results.
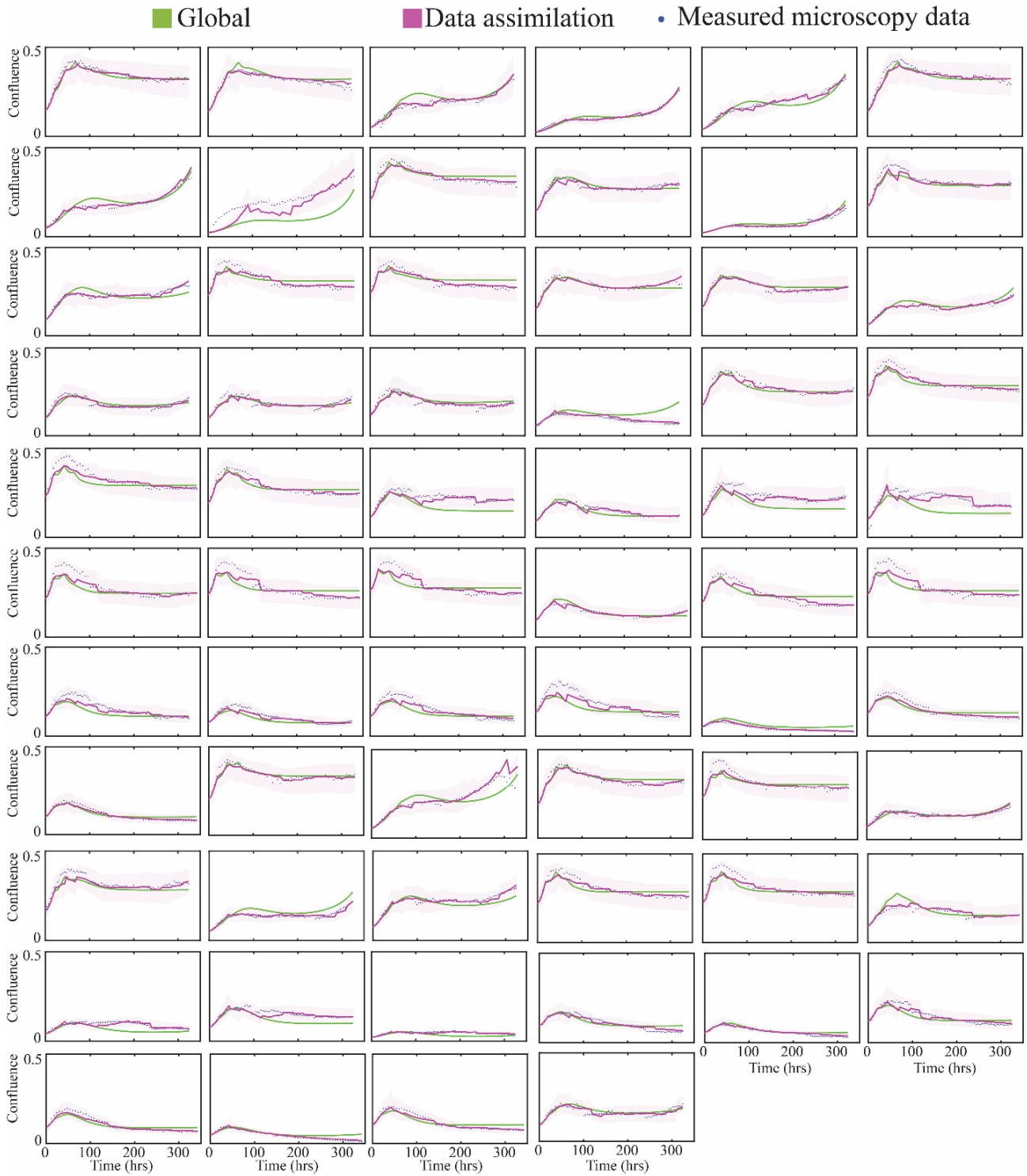
Figure S4. The data assimilation prediction versus global prediction for all C6 samples from the validation group. See caption to Figure 6 for an explanation of the results.

Table S1. Model parameters.

Parameter	Units	Interpretation	Source
k_p	hr^{-1}	Proliferation rate	Computed from
θ	1	Carrying capacity	control group, and
A	1	Allee effect	fixed throughout experiments
N_p	1	Confluence of proliferating cells	Initial cell confluence and total
N_s	1	Confluence of senescent cells	confluence (i.e.,
N_0	1	Initial confluence of cells at time $= 0$	N_p+N_s) are measured from microscopy data.
f_{DSB}	1	The fraction of DSBs remaining unrepaired (normalized between 0 and 1)	Measured by flow cytometry
k_{acute}	1	Early death rate	Fit for individual
k_{accum}	1	Late death rate	parameters $X_{ind,i}$
$k_{acute,N}$	hr^{-1}	Death rate quantifying the contribution of initial confluence to early death	Fit for global parameters X_{pop}
$a_{accum,N}$	1	Scale factor quantifying the contribution of initial confluence to late death	
$k_{accum,D}$	hr^{-1}	Death rate quantifying the contribution of radiation doses to late death	
r	hr^{-1}	Radiation efficacy	Fit both globally and
k_{ps}	hr^{-1}	Conversion rate from proliferation to senescent components	individually

Note: A units value of "1" indicates the parameter is dimensionless.

Table S2. Pre-trained cell constants (mean ± standard error).

Parameter	9L	C6
k_p	0.079 (+/- 0.011)	0.16 (+/- 0.019)
θ	0.98 (+/- 0.013)	0.81 (+/- 0.0061)
A	0.33 (+/- 0.043)	0.15 (+/- 0.0076)

S7. 9L validation group versus testing group accuracy.

To understand why the accuracy of 9L validation group (0.953 ± 0.052) is better compared to its testing group (0.950 ± 0.049), which is not common in learning models, we checked the accuracy of 9L validation group and testing group for each treatment conditions, and acquired the following numbers (mean ± standard deviation):

a) Four fractions of 4 Gy – testing (0.961 ± 0.027), validation (0.960 ± 0.029).
b) Three fractions of 5.3 Gy – testing (0.878 ± 0.050), validation (0.875 ± 0.092).
c) Two fractions of 8 Gy – testing (0.916 ± 0.041), validation (0.942 ± 0.036).
d) Four fractions of 5 Gy – testing (0.984 ± 0.004), validation (0.984 ± 0.002).
e) Three fractions of 6.7 Gy – testing (0.986 ± 0.005), validation (0.986 ± 0.005).
f) Two fractions of 10 Gy – testing (0.977 ± 0.013), validation (0.974 ± 0.010).

This confirms that the slight difference only comes from the two fractions of 8 Gy group of the 9L cell lines. We then checked the original data and there were no curious differences found. However, we do notice the following: because the testing group only includes 36 9L samples in total as mentioned in the Methods section (and only 6 samples are from two fractions of 8 Gy group), a 'bad prediction' curve may greatly affect this value. In fact, there are two such curves in these six samples (with CCCs of only 0.857 and 0.891, respectively). By removing these two 'bad examples', the average of the remaining four curves achieves 0.947 ± 0.027. We thus believe this slight difference is due to the two outliers and small sample size used in our testing set.