Online Appendix

Section A: Weighted average visit time

A weighted average is used when the components being averaged are not equally represented. Obviously, in clinical practice, the number of new versus established patients will vary, as will their complexity. The Centers for Medicare & Medicaid Services (CMS) has provided data on how frequently physicians submit claims for each E/M CPT code, with new versus established codes reported separately (https://www.cms.gov/data-research/statistics-trends-and-reports/medicare-fee-for-service-parts-a-b/medicare-utilization-part-b.)

For the accompanying editorial (https://www.aafp.org/fpm/2026/0700/sustainable-workload), I needed to calculate the average predicted time spent per visit across all E/M office visit levels. To give the most conservative estimate, I used the minimum time allotted per code and assumed 6 new patients and 94 established patients per 100 visits, also based on the CMS data. I then multiplied the percentage of each new patient code by 0.06 and each established patient code by 0.94. Next, I multiplied the minimum time allotted for each code by its calculated weight. Finally, I summed the weighted times for individual codes to provide the weighted average time per visit — 26 minutes.

Table A: Weighted average calculation of time spent per patient

	CPT	Minimum time allotted per CPT (min)	Percent billed per CMS	Weight (0.06 for new patients, 0.94 for established patients)	Weighted time (min)
New	99201	0	1.4%	0.00084	0
	99202	15	13.5%	0.0081	0.1215
	99203	30	47.7%	0.02862	0.8586
	99204	45	33.6%	0.02016	0.9072
	99205	60	3.8%	0.00228	0.1368
Established	99211	0	2.4%	0.02256	0
	99212	10	2.3%	0.02162	0.2162
	99213	20	37.4%	0.35156	7.0312
	99214	30	54.9%	0.51606	15.4818
	99215	40	3.1%	0.02914	1.1656
	Weighted average visit time (min)				25.9189

Section B: Line queuing model

The M/M/1 line queuing model uses two inputs: service time and appointment length. Using those variables, mean arrival rate and mean service rate can be calculated as follows:

λ = Mean arrival rate = 1/appointment length

μ = Mean service rate = 1/service time

Using these two factors, we can calculate the other outputs:

Physician time utilization = ρ = λ/ μ

Mean number of patients in office = L = $\frac{\rho}{1-\rho}$

Mean number of patients waiting = Lq = $\frac{\rho^2}{1-\rho}$

Mean time patient is in office = W = L/λ

Mean time patient waits = Wq = Lq/λ

Table B: Line queuing model

	Clinic A		Clinic B	Physician time utilization set to 80%							
Mean appt length (min)	15	15	25.5	34	25.5	20.4	17	14.571	12.75	11.333	10.2
Mean service time (min)	25.9	12	20	27.2	20.4	16.32	13.6	11.657	10.2	9.067	8.16
Mean arrival rate	0.067	0.067	0.039	0.029	0.039	0.049	0.059	0.069	0.078	0.088	0.098
Mean service rate	0.039	0.083	0.05	0.037	0.049	0.061	0.074	0.086	0.098	0.11	0.123
Mean number of pts in office	Error	4	3.636	4	4	4	4	4	4	4	4
Mean number of pts waiting	Error	3.2	2.852	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2
Mean time pt spends in office (min)	Error	60	92.727	136	102	81.6	68	58.286	51	45.333	40.8
Mean time pt waits (min)	Error	48	72.727	108.8	81.6	65.28	54.4	46.629	40.8	36.267	32.64
Physician time utilization	1.727	0.8	0.784	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Extra time at end of 8.5-hr workday (min)	Error	404.6	28.	-106.5	28	162.5	297	431.5	566	700.5	835
Number of pts per 8.5-hr workday	32	33	34	15	20	25	30	35	40	45	50

Main article: Trapp C. Estimating a sustainable workload in primary care: implications on the physician shortage. *Fam Pract Manag*. 2026;33;4:7-9.