

**Supplemental Table 1: Number of Physicians Identified, by Specialty, Year, and Dataset**

Specialty	2014				2015				2016			
	Procedures	%	Prescriptions	%	Procedures	%	Prescriptions	%	Procedures	%	Prescriptions	%
Allergy / Immunology	2,850	0.5%	3,493	0.7%	2,895	0.5%	3,594	0.7%	2,922	0.5%	3,647	0.7%
Anesthesiology	35,088	6.7%	7,707	1.5%	35,640	6.7%	7,744	1.4%	36,299	6.7%	7,827	1.5%
Cardiology	22,510	4.3%	23,602	4.5%	22,786	4.3%	23,831	4.5%	21,451	4.0%	24,252	4.5%
Dermatology	9,776	1.9%	11,516	2.2%	9,960	1.9%	11,763	2.2%	10,150	1.9%	12,004	2.2%
Emergency Medicine	38,304	7.3%	35,746	6.8%	39,273	7.4%	36,561	6.8%	40,285	7.5%	36,281	6.7%
Endocrinology	4,809	0.9%	5,272	1.0%	4,963	0.9%	5,390	1.0%	5,123	1.0%	5,585	1.0%
Gastroenterology	11,534	2.2%	12,463	2.4%	11,747	2.2%	12,615	2.4%	12,010	2.2%	12,849	2.4%
Hematology-Oncology	11,307	2.2%	11,541	2.2%	11,623	2.2%	11,828	2.2%	11,896	2.2%	12,026	2.2%
Infectious Disease	4,863	0.9%	4,405	0.8%	4,998	0.9%	4,559	0.9%	5,155	1.0%	4,682	0.9%
Nephrology	7,282	1.4%	7,803	1.5%	7,496	1.4%	7,997	1.5%	7,715	1.4%	8,172	1.5%
Neurology	12,081	2.3%	12,599	2.4%	12,427	2.3%	12,836	2.4%	12,852	2.4%	13,074	2.4%
Neurosurgery	3,939	0.8%	3,530	0.7%	4,027	0.8%	3,471	0.6%	4,084	0.8%	3,428	0.6%
OB/GYN	20,816	4.0%	24,736	4.7%	20,540	3.9%	24,492	4.6%	20,416	3.8%	24,090	4.5%
Ophthalmology	15,997	3.0%	18,122	3.4%	16,072	3.0%	18,240	3.4%	16,199	3.0%	18,345	3.4%
Orthopedic Surgery	20,810	4.0%	22,349	4.2%	21,222	4.0%	22,423	4.2%	21,524	4.0%	22,554	4.2%
Otolaryngology	7,561	1.4%	8,982	1.7%	7,624	1.4%	9,056	1.7%	7,648	1.4%	9,090	1.7%
PM&R	6,710	1.3%	6,779	1.3%	6,856	1.3%	6,963	1.3%	6,969	1.3%	7,039	1.3%
Pathology	10,680	2.0%	455	0.1%	10,710	2.0%	412	0.1%	10,787	2.0%	375	0.1%
Plastic Surgery	3,277	0.6%	3,675	0.7%	3,292	0.6%	3,663	0.7%	3,305	0.6%	3,607	0.7%
Primary Care	177,498	33.8%	215,512	40.8%	179,805	33.8%	218,705	40.9%	181,762	33.7%	220,773	41.0%
Psychiatry	20,017	3.8%	34,440	6.5%	20,026	3.8%	34,683	6.5%	20,016	3.7%	34,681	6.4%
Pulmonology	11,150	2.1%	9,408	1.8%	11,604	2.2%	9,568	1.8%	12,033	2.2%	9,708	1.8%
Radiation Oncology	3,822	0.7%	2,951	0.6%	3,912	0.7%	2,973	0.6%	3,967	0.7%	3,097	0.6%
Radiology	25,410	4.8%	2,281	0.4%	25,757	4.8%	2,272	0.4%	26,246	4.9%	2,215	0.4%
Rheumatology	3,936	0.7%	4,362	0.8%	4,076	0.8%	4,482	0.8%	4,161	0.8%	4,595	0.9%
Surgery	25,040	4.8%	25,119	4.8%	25,400	4.8%	25,245	4.7%	25,756	4.8%	25,033	4.6%
Urology	7,899	1.5%	9,867	1.9%	7,952	1.5%	9,935	1.9%	7,973	1.5%	9,897	1.8%
<b>Total</b>	<b>524,966</b>		<b>528,715</b>		<b>532,683</b>		<b>535,301</b>		<b>538,704</b>		<b>538,926</b>	

Supplemental materials for: Wingrove P, Liaw W, Weiss J, Petterson S, Maier J, Bazemore A. Using machine learning to predict primary care and advance workforce research. *Ann Fam Med.* 2020;18(4):334-340.

Notes: Procedure data are from the 2014-2016 Centers for Medicare & Medicaid Services (CMS) Provider Utilization & Payment Data Public Use File. Prescribing data are from the 2014-2016 CMS Part D Prescriber Public Use File.

OB/GYN = Obstetrics and Gynecology

PM&R = Physical Medicine & Rehabilitation

**Supplemental Table 2: Mean Combined Model Recall, Precision, and Negative Predictive Value with 95% Confidence Intervals by Specialty**

Specialty	Recall (Sensitivity)	Precision (PPV)	NPV
Allergy / Immunology	.902[.895 - .909]	.922[.914 - .929]	.999[.999 - .999]
Anesthesiology	.95[.949 - .952]	.951[.949 - .953]	.997[.997 - .997]
Cardiology	.922[.918 - .925]	.954[.952 - .956]	.997[.997 - .997]
Dermatology	.972[.97 - .973]	.96[.957 - .963]	.999[.999 - .999]
Emergency Medicine	.878[.875 - .881]	.916[.914 - .918]	.991[.991 - .992]
Endocrinology	.836[.822 - .849]	.896[.891 - .901]	.999[.998 - .999]
Gastroenterology	.931[.927 - .936]	.915[.912 - .919]	.999[.998 - .999]
Hematology-Oncology	.866[.861 - .871]	.883[.879 - .886]	.997[.997 - .997]
Infectious Disease	.645[.638 - .652]	.882[.877 - .887]	.997[.997 - .997]
Nephrology	.886[.883 - .889]	.883[.881 - .885]	.998[.998 - .999]
Neurology	.852[.848 - .856]	.921[.916 - .926]	.997[.997 - .997]
Neurosurgery	.543[.533 - .552]	.812[.801 - .824]	.997[.997 - .997]
OB/GYN	.917[.915 - .918]	.923[.917 - .929]	.996[.996 - .997]
Ophthalmology	.982[.982 - .983]	.982[.981 - .983]	.999[.999 - .999]
Orthopedic Surgery	.881[.876 - .886]	.888[.882 - .893]	.995[.995 - .995]
Otolaryngology	.906[.901 - .91]	.96[.955 - .964]	.999[.998 - .999]
PM&R	.477[.474 - .479]	.761[.756 - .766]	.994[.994 - .994]
Pathology	.982[.981 - .984]	.991[.99 - .991]	1[1 - 1]
Plastic Surgery	.477[.471 - .483]	.604[.592 - .616]	.997[.997 - .997]
Primary Care	.936[.935 - .938]	.903[.901 - .905]	.964[.963 - .965]
Psychiatry	.929[.928 - .929]	.932[.93 - .935]	.996[.996 - .996]
Pulmonology	.846[.843 - .849]	.826[.824 - .828]	.997[.997 - .997]
Radiation Oncology	.934[.928 - .939]	.945[.934 - .956]	1[1 - 1]
Radiology	.968[.967 - .968]	.99[.989 - .991]	.999[.999 - .999]
Rheumatology	.928[.923 - .934]	.903[.899 - .908]	.999[.999 - 1]
Surgery	.822[.813 - .831]	.731[.716 - .747]	.991[.991 - .992]
Urology	.953[.948 - .958]	.97[.964 - .976]	.999[.999 - .999]

Supplemental materials for: Wingrove P, Liaw W, Weiss J, Petterson S, Maier J, Bazemore A. Using machine learning to predict primary care and advance workforce research. *Ann Fam Med.* 2020;18(4):334-340.

OB/GYN = Obstetrics and Gynecology

PM&R = Physical Medicine & Rehabilitation

PPV = Positive Predictive Value

NPV = Negative Predictive Value

**Supplemental Table 3. Area under the Receiver Operating Characteristic (ROC) Curve, by Specialty, for the 2016 Combined Random Forest Applied to the 2016 Test Data**

<b>Specialty</b>	<b>Area Under the ROC Curve</b>
Allergy / Immunology	0.995
Anesthesiology	0.997
Cardiology	0.997
Dermatology	0.998
Emergency Medicine	0.985
Endocrinology	0.995
Gastroenterology	0.998
Hematology-Oncology	0.991
Infectious Disease	0.992
Nephrology	0.992
Neurology	0.994
Neurosurgery	0.988
OB/GYN	0.995
Ophthalmology	0.999
Orthopedic Surgery	0.995
Otolaryngology	0.994
PM&R	0.998
Pathology	0.991
Plastic Surgery	0.984
Primary Care	0.982
Psychiatry	0.997
Pulmonology	0.992
Radiation Oncology	0.992
Radiology	0.998
Rheumatology	0.996

Supplemental materials for: Wingrove P, Liaw W, Weiss J, Petterson S, Maier J, Bazemore A. Using machine learning to predict primary care and advance workforce research. *Ann Fam Med.* 2020;18(4):334-340.

Surgery	0.984
Urology	0.998

---

**Note:**

ROC = Receiver Operating Characteristic

OB/GYN = Obstetrics and Gynecology

PM&R = Physical Medicine & Rehabilitation