Appendix 1: Physician workload in the Timmins FHT

Statistical Methods

1. Roster size corrections

Nursing supports — Some members of the Timmins FHT have greater degrees of nurse-led care in their offices. This variation may might affect the total number of patients a GP can support on his or her roster.

To correct for relative differences in nursing supports among clinic sites, we obtained the number of full-time-equivalent nurses (excluding nurse practitioners) at a given site, divided by the total number of GPs working at that site. We assumed these nursing support values to be normally distributed and standardized these values (i.e., $\frac{x_i - \bar{x}}{\sigma}$) for each site. Next, we assumed that increasing or decreasing levels of nursing support can support panel sizes about 10% larger or smaller at the greatest and lowest levels of support, respectively (Muldoon et al. 2012). By definition, the highest level of nursing support is about 1.96 standard deviations above the mean (i.e. the 97.5% upper confidence interval of a standard normal distribution). We therefore multiplied the standardized values by $\frac{10\%}{1.96}$, or about 5.1%, and then adjusted each corresponding roster size to inflate or deflate them accordingly. For instance, if GP A at clinic X_1 had nursing support one standard deviation above the Timmins mean, whereas GP B at clinic X_2 had only the mean level of support, we deflated panel size of GP A to 94.9% of its initial value, effectively decreasing the number of patients GP A can support without his or her enhanced nursing compared to GP B, whose patient roster size remains unchanged. This assumes an equal proportion of nurse-led visits among all patients for a given GP.

Integrating the corrections — We compared each of these "practical" roster size correction magnitudes against each other, then also calculated combined corrected roster sizes by applying each percentage change in sequence.

We did not correct for the presence or absence of learners, as their attendance is not consistent through the year or among offices and preceptors. Residents and learners cycle between all of the locations of the Timmins FHT throughout the year. Thus, we assumed the learner effect to be consistent across sites, and ultimately a consistent influence on patient throughput.

For the remainder of the analysis, we used the corrected roster size.

2. GP availability (γ_{cor})

Availability and continuity for a given roster were assessed by calculating the average number of patient visits per year ("γ") as a relevant patient-oriented outcome.

Methodologically, γ_{cor}, the corrected availability, was calculated by dividing the number of appointments per year for each GP divided by his or her corrected roster sizes (see above), as in:

$$\gamma_{cor} = \frac{totalAppointments}{rosterSize} \quad (1)$$

We compared these values among Timmins FHT GPs by presenting histograms of the aggregated data.

To indirectly test for decreased GP availability in larger rosters, we modeled the second-order relationship, if any, between the formal and practical roster sizes calculated above. A negative relationship would imply a proportionate increase in dormant or unseen patients with increasing formal roster size.

With these data, we also built a linear model to test whether GP availability (γ_{cor}) varied with corrected roster size, effectively modeling the variation in total annual appointments across the range of roster sizes (see Equation 1 for illustration).

Values of γ_{cor} were then used in the models of patient diversion described below.

3. Chronic disease prevalence and availability

We used linear models to test whether GP availability (γ) decreased with higher chronic disease burdens (calculated above) among rosters. We compared the effect of the proportion of each of CHF, COPD and diabetes, both with and without a linear correction term for corrected roster size.

4. Patient diversions from their rostered GP

Diversion to the Emergency Department for daytime, low-acuity visits

For ED diversion, we considered low-acuity visits during working hours as the most-representative of "replacement" care. This recognizes that higher-acuity events, and events perceived as necessary outside of working hours may not have an alternative primary care option. In this case, we modeled the relationship between the number of low-acuity (CTAS 4 or 5), daytime (8:00 am to 5:00 pm) ED visits and our corrected measure of GP availability (visits per patient per year; γ_{cor}).

The model framework described $y_{1,i}$, the number of daytime, low-acuity ED visits per patient per year, as a negative binomial distribution to best account for the zero-inflation in the sample:

$$y_{1,i} \sim negBinom(e^{X_i\beta}) \quad (2)$$

where $y_{1,i}$ is the count of ED visits for patient $i = 1, \ldots, n$, and $X_i\beta$ describes the linear terms:

$$X_i\beta = \alpha + \beta_\gamma \cdot \gamma_{i[j]} + (\beta_{DM} \cdot x_{DM \cdot i}) + (\beta_{COPD} \cdot x_{COPD \cdot i}) + (\beta_{CHF} \cdot x_{CHF \cdot i}) \\ +(\beta_{male} \cdot x_{male \cdot i}) + (\beta_{infant} \cdot x_{infant \cdot i}) + (\beta_{child} \cdot x_{child \cdot i}) + (\beta_{senior} \cdot x_{senior \cdot i}) \quad (3)$$

where γ refers to γ_{cor}, and the corresponding sex, age-category and chronic disease indicators ($x_{CHF \cdot i}$, $x_{COPD \cdot i}$, $x_{DM \cdot i}$) have null conditions of female sex, adult (age 18-65) status, and the absence of each chronic disease. Infants were under age 2. The variance of $X_i\beta$ is equal to itself multiplied by an internally-estimated dispersion parameter θ.

We did not ascribe random intercepts for clinician $j = 1, \ldots, m$ explicitly as they would be largely confounded with the unique values of $\gamma_{i[j]}$ for each clinician.

Our principal value of interest was β_γ, the effect of GP availability on the number of ED attendances through the year.

To best assess for the possibility of threshold values of γ_{cor} above which diversion rates markedly increase, we did not model β_γ as a single, linear coefficient. Rather, we built a generalized additive model ("GAM"; Wood 2006) in place of the linear model. Compared to generalized linear models, GAMs use smoothing techniques to estimate model coefficients as entire nonlinear functions, rather than as single values. By visualizing these functions across the range of the independent variable, GAMs can appreciate underlying patterns that are invisible to first-order coefficients alone. This approach is similar to the cubic splines used by Dahrouge et al (2016) in their models of patient outcomes. Here, we visualized β_γ as a nonlinear function to test whether its effect changed with different values of γ_{cor}, and whether a threshold value might exist. We expected to see lower values (corresponding to decreased ED visits) as γ_{cor} gets larger.

Hence, the actual negative binomial model terms tested were:

$$X_i\beta = \alpha + f_1(\beta_\gamma) \cdot \gamma_{i[j]} + (\beta_{DM} \cdot x_{DM \cdot i}) + (\beta_{COPD} \cdot x_{COPD \cdot i}) + (\beta_{CHF} \cdot x_{CHF \cdot i}) \\ + (\beta_{male} \cdot x_{male \cdot i}) + (\beta_{infant} \cdot x_{infant \cdot i}) + (\beta_{child} \cdot x_{child \cdot i}) + (\beta_{senior} \cdot x_{senior \cdot i}) \quad (4)$$

where $f_1()$ is the smoothing function acting on β_γ in this case.

We enforced a maximum order of the smoothing function of 3, since we did not expect the effect of γ_{cor} to vary periodically or at higher-order patterns.

Diversion to walk-in clinics

Asking an essentially identical question as above, we also modeled the count of after-hours walk-in clinic visits ($y_{2,i}$) as the dependent variable in place of the count of daytime, low-acuity ED visits per patient per year ($y_{1,i}$). Again, we used a GAM to test for nonlinear effects of GP availability (γ_{cor}) on the outcome.

Model diagnostics

Model diagnostics included QQ plots and posterior probability plots from Bayesian equivalents (Stan Development Team 2016). In the model of WIC visits, these suggested a good fit to the data. The model also explained a larger proportion of deviance (2.9%) compared to the ED model (<1%).

References

1. Green LV, Savin S, Murray M. Providing timely access to care: what is the right patient panel size? *Jt Comm J Qual Patient Saf*. 2007;33(4):211-218.
2. Muldoon L, Dahrouge S, Russell G, Ward WH and N. How many patients should a family physician have? Factors to consider in answering a deceptively simple question. Healthcare Policy. http://www.longwoods.com/content/22885. Accessed May 26, 2017.

3. Murray M, Davies M, Boushon B. Panel Size: How Many Patients Can One Doctor Manage? *FPM*. 2007;14(4):44. http://www.aafp.org/fpm/2007/0400/p44.html. Accessed May 28, 2017.
4. Stan Development Team. *Rstanarm: Bayesian Applied Regression Modeling via Stan*.; 2016. http://mc-stan.org/.
5. Wood, Simon N. *Generalized Additive Models: An Introduction with R*. New York: Chapman and Hall; 2006.