Modelling patient flow and outcomes in community healthcare – a fluid approximation of a stochastic queueing system

Ryan Palmer, Martin Utley

Clinical Operational Research Unit (CORU),
University College London (UCL), London
e-mail: ryan.palmer.14@ucl.ac.uk

IFORS 2017 - Québec City

Context: Community healthcare and healthcare policy

In recent decades, an ambition of healthcare policy has been to deliver more care in the community sector [6].

- Diverse range of services, operating in different physical locations
- Common for patients to use a range of services which they may re-use
- Considered to be crucial in meeting the current and future challenges that face modern health care services [3]

Challenge: how to organise and deliver these services given: physical distribution, patients using multiple services, increased referrals, case mix, and long term care requirements [7].

Quality measurement in community healthcare

- Operational capability - waiting time, queue length, resource utilisation, capacity
- Outcome measures - aspects of a patient's health or experience, influenced by a care interaction
  - i.e. measurable behaviours, opinions, medical characteristics or health status
  - Used to monitor and evaluate the progression of patients and the quality of care received

Increasingly used by managers, clinicians and commissioners to inform quality improvement [2].

Outcomes in practice and modelling

Uses in healthcare:

- Regularly evaluated as periodically calculated proportions
  - Misses time dependent variability in output of service and outcomes achieved
- Often considered in isolation from other services
- Misleading when considering dynamic patient flow of stochastic healthcare systems

Common modelling assumptions:

- Operational improvements positively affect outcomes
- Uniform patients

Rationale and aims

Understand dynamics of patient flow and use of patient outcomes in evaluating community healthcare

1. Unify two perspectives of quality in a single modelling framework
   - Account for flow of patients in different health states
   - Transition in health throughout queueing process important to understanding demand for and effect of system
2. Establish a concept of the flow of outcomes - how individual services contribute to a system's performance
   - Flow: bottlenecks, required capacity, waiting times
   - Outcomes: how they accrue over time through a combination of services

Dynamics of patient flow in community healthcare

- Multiple services of varying configuration
- Patients reusing the same services - either sequentially or after care within another service
- Possibility of patients abandoning queue - impatient; seek healthcare in a non-community setting
- Patients arrive in different health states
  - Different capacity to benefit/resource requirement
  - Health may improve or decline throughout
- Time dependent demand

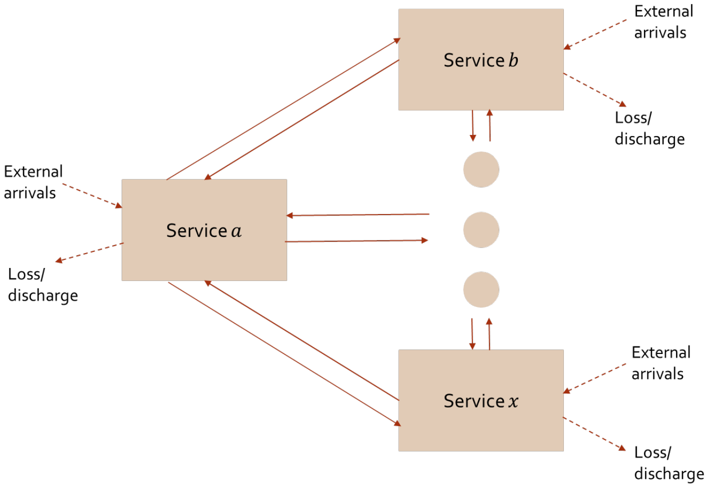
Traditional methods do not cope well with these dynamics - computationally expensive

Extension and application of a fluid approximation

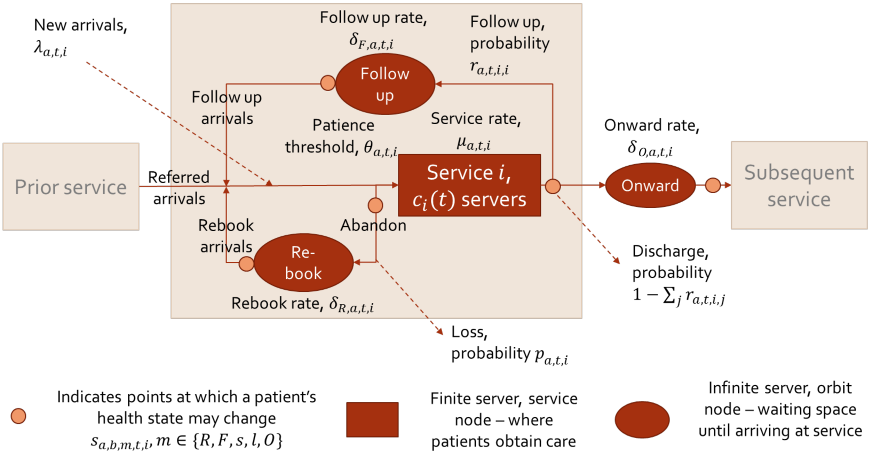
We present the application of a deterministic fluid and diffusion approximation for a stochastic queueing system. A novel application and extension of work by both S Ding et al. (2015)[1] and A Mandelbaum et al. (1998, 2002) [4, 5].

- Network of multiple services
- Health states - different parameters
- Application of diffusion equation

The system - a series of stochastic processes



The system - a series of stochastic processes



Number of servers split across A parallel queues

Fluid approximation of patient flow

For analytical tractability, patients served first come first serve in parallel queues according to health state, $a \in A$.

Capacity $C_{a,i}(t)$ allocated to each queue at time t, with $\sum_{a \in A} C_{a,i}(t) = c_i(t)$.

$$C_{a,i}(u, \mathbf{Z}(u)) = \begin{cases} \frac{c_i Z_{a,Q,i}(u)}{\sum_{b=1}^{A} Z_{b,Q,i}(u)}, & \text{if } c_i < \sum_{b=1}^{A} Z_{b,Q,i}(u) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$C_{a,i}(u, \mathbf{Z}(u)) = \begin{cases} \frac{c_i \mu_{a,i} Z_{a,Q,i}(u)}{\sum_{b=1}^{A} \mu_b Z_{b,Q,i}(u)}, & \text{if } c_i < \sum_{b=1}^{A} Z_{b,Q,i}(u) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$C_{a,i}(u, \mathbf{Z}(u)) = \frac{Z_{a,Q,i}(0) + \lambda_{a,i}}{\sum_{b=1}^{A} Z_{b,Q,i}(0) + \lambda_{b,i}} \times c_i \quad (3)$$

Fluid approximation of patient flow

- Work within Skorokhod space with J1 metric: intuitively provides wiggle space within both space and time.
- A natural and convenient formalism for describing trajectories of stochastic processes that may admit discontinuities, i.e. trajectories of Poisson processes.

Consider a sequence of models where the n-th model denoted by the superscript (n) has an arrival rate of $\lambda_{a,i} n$ for new patients in health state a and the total number of servers is nc_i. The scaled fluid process is defined as:

$$\bar{Z}_{a,m,i}(t) := \frac{Z^{(n)}_{a,m,i}(t)}{n},$$

where $a \in \{1, ..., A\}$, $i = 1, ..., J$ and $m \in \{Q, R, F, O, D, L\}$

Within this space and scaling the system for $n \to \infty$, we can define the fluid limit:

$$
\begin{aligned}
z_{a,Q,i}(t) = z_{a,Q,i}(0) + \lambda_{a,i} t &+ \sum_{b=1}^{A} s_{b,a,R,i}\ \delta_{b,R,i} \int_0^t z_{b,R,i}(u)\ du \\
&+ \sum_{b=1}^{A} s_{b,a,F,i}\ \delta_{b,F,i} \int_0^t z_{b,F,i}(u)\ du \\
&+ \sum_{b=1}^{A} s_{b,a,O,i}\ \delta_{b,O,i} \int_0^t z_{b,O,i}(u)\ du \\
&- \mu_{a,i} \int_0^t \min\left(z_{a,Q,i}(u), C_{a,i}(u, \mathbf{z}(u))\right)\ du \\
&- \theta_{a,i} \int_0^t \left(z_{a,Q,i}(u) - C_{a,i}(u, \mathbf{z}(u))\right)^+\ du
\end{aligned}
$$

Fluid approximation of patient flow

$$
\begin{aligned}
z_{a,R,i}(t) &= z_{a,R,i}(0) - \delta_{a,R,i} \int_0^t z_{a,R,i}(u)\ du \\
&\quad + p_{a,i} \sum_{b=1}^{A} s_{b,a,l,i}\ \theta_{b,i} \int_0^t \left(z_{b,Q,i}(u) - C_{b,i}\left(u, \mathbf{z}(u)\right)\right)^+\ du \\
z_{a,F,i}(t) &= z_{a,F,i}(0) - \delta_{a,F,i} \int_0^t z_{a,F,i}(u)\ du \\
&\quad + r_{a,i,i} \sum_{b=1}^{A} s_{b,a,s,i} \int_0^t \mu_{b,i} \min\left(z_{b,Q,i}(u), C_{b,i}(u, \mathbf{z}(u))\right)\ du
\end{aligned}
$$

Fluid approximation of patient flow

$$z_{a,O,i}(t) = z_{a,O,i}(0) - \delta_{a,O,i} \int_0^t z_{a,O,i}(u)\ du$$
$$+ \sum_{j=1;\ j\neq i}^{J} \sum_{b=1}^{A} s_{b,a,s,j} r_{a,j,i} \int_0^t \mu_{b,j} \min\left(z_{b,Q,j}(u), C_{b,j}(u, \mathbf{z}(u))\right)\ du$$

$$z_{a,L,i}(t) = z_{a,L,i}(0)$$
$$+ (1 - p_{a,i}) \sum_{b=1}^{A} s_{b,a,l,i}\ \theta_{b,i} \int_0^t \left(z_{b,Q,i}(u) - C_{b,i}\left(u, \mathbf{z}(u)\right)\right)^+\ du$$

$$z_{a,D,i}(t) = z_{a,D,i}(0) + r_{a,i,D} \sum_{b=1}^{A} s_{b,a,s,i}\ \mu_{b,i} \int_0^t C_{b,i}(u, \mathbf{z}(u))\ du$$

Fluid approximation of patient flow

Analytical expressions cannot be found for the above, however can be solved iteratively using common numerical schemes.

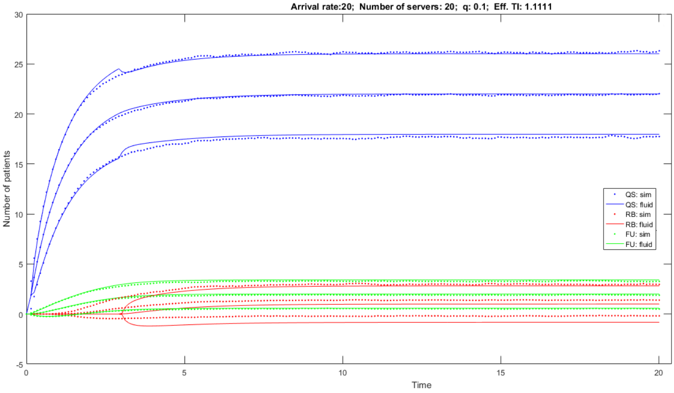
- Rewrite $z_{a,Q,i}(t), z_{a,R,i}(t), z_{a,F,i}(t), z_{a,O,i}(t),\ \ i = 1, ..., J,\ \ a = 1, ..., A$ as $z(t) = \phi(z(t))$
- Let $z^{(0)}(0) = 0$
- Calculate $z^{(k+1)} = \phi(z^{(k)}), k = 0, 1, ...$ using a common numerical scheme
- Stop when difference between $z^{(k+1)}$ and $z^{(k)}$ is deemed sufficiently small

Computed in MATLAB, produced a Discrete Event Simulation for the basic stochastic system and extension with health states

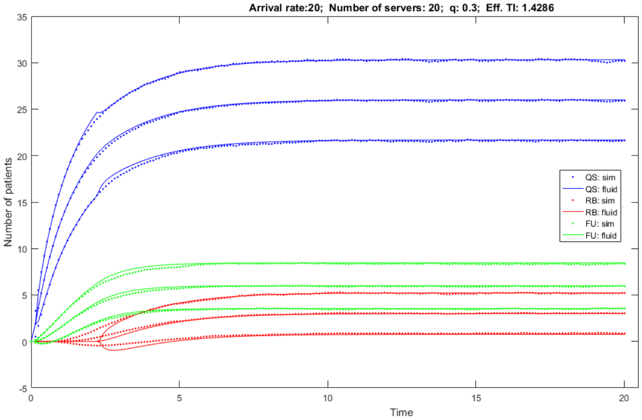
- Compare fluid model to simulation
- Basic model - rebook and follow up [1]
  - Explore parameter space for community healthcare
  - Triangulation of models
- Extended to include health states - single service
  - Assess accuracy of extension

Simulations - Basic case
Small system with small probability of re-use

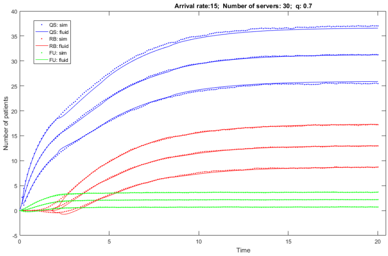
Effective traffic intensity: $\hat{\rho} = \frac{\lambda}{c\mu(1-q)}$

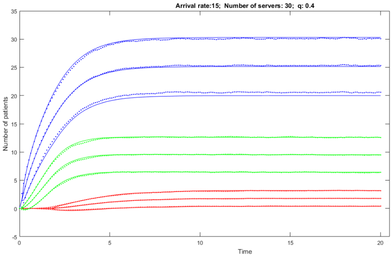


Simulations - Basic case
Larger probability of re-use

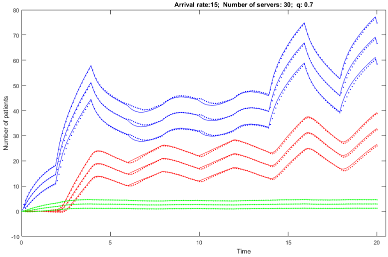


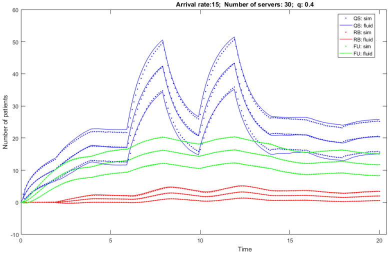
Simulations - Introduction of outcomes





Simulations - System with outcomes
Time dependent behaviour - arrival spikes





Measures from fluid approximation of patient flow

- Number of patients, over time, in:
  - Services (or orbits)
  - Health states
- Waiting times, number in the system, waiting time distribution - using Erlang A or R, Virtual Waiting Time [5]
  - Per service
- Production of outcomes: number of patients discharge from a service/system in a given health state over time
- Loss over time

Types of analyses

- Capacity allocation - i.e. optimisation:
  1. Balance across queues for equitable wait times
  2. Reduce net loss
  3. Prevent resource intensive re-joins
- Production of outcomes:
  1. Begin to understand how a network of services work together to "produce" patients with good health
  2. Seek balance across multiple services
     - Identify bottlenecks
     - Holistic view of operational measures

Limitations and future work

Limitations:

- A deterministic analogue of a stochastic system
- Errors for none heavily loaded systems
- Less accurate for smaller systems

Future directions:

1. Joint use of services - could an extension capture this?
2. Can the patient flow and outcomes of patients with multiple morbidities be informatively modelled?
3. Combining with optimisation methods - can useful information be gained for service planning?

Conclusions

Extending a fluid approximation of stochastic systems to a network of services, including patient health, is beneficial since:

- Model key dynamics of community healthcare
- Overcomes computational burden and time expense of other methods
- Provides time dependent analysis of system outputs

References

[1] Ding S, Remerova M, van der Mei RD, and Zwart B.
Fluid approximation of a call center model with redials and reconnects.
Performance Evaluation, 92:24–39, 2015.

[2] Foot C, Sonola L, Bennett L, Fitzsimons B, Raleigh V, and Gregory S.
Managing quality in community health care services.
London: The King's Fund, 2014.

[3] Ham C, Dixon A, and Brooke B.
Transforming the delivery of health and social care: the case for fundamental change.
London: Kings Fund, 2012.

[4] Mandelbaum A, Massey W A, and Reiman M I.
Strong approximations for markovian service networks.
Queueing Systems, 30(1):149–201, 1998.

[5] Mandelbaum A, Massey W A, Reiman M I, Stolyar A, and Rider B.
Queue lengths and waiting times for multiserver queues with abandonment and retrials.
Telecommunication Systems, 21(2):149–171, 2002.

[6] Munton T, Martin A, Marrero I, Llewellyn A, Gibson K, and Gomersall A.
Evidence: Getting out of hospital?
2011.

[7] Palmer R, Fulop N J, and Utley M.
A systematic literature review of operational research methods for modelling patient flow and outcomes within community healthcare and other settings.
Health Systems, pages 1–21, 2017.

Acknowledgements

This research is supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care North Thames at Barts Health NHS Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The author is supported by the Health Foundation as part of the Improvement Science PhD programme. The Health Foundation is an independent charity committed to bringing about better health and healthcare for people in the UK.

The Health Foundation
Inspiring Improvement

Collaboration for Leadership in Applied Health Research and Care North Thames

http://www.clahrc-norththames.nihr.ac.uk

Funded by NIHR

Thank you for your attention

Are there any questions?