Maximising access to thrombectomy services for stroke in England: a modelling study.

Michael Allen[1], Kerry Pearn[1], Martin James[2], Gary A. Ford[3], Phil White[4], Anthony G. Rudd[5], Peter McMeekin[6], Ken Stein[1]

[1]University of Exeter Medical School and the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) South West Peninsula.

[2]Royal Devon and Exeter NHS Foundation Trust and the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) South West Peninsula

[3]Radcliffe Department of Medicine, Oxford University and Oxford University Hospitals NHS Foundation Trust

[4]Institute of Neuroscience, Newcastle University and Newcastle upon Tyne Hospitals NHS Foundation Trust

[5]Kings College London and Guy's and St Thomas, NHS Foundation Trust.

[6]Faculty of Health and Life Sciences, Northumbria University

Corresponding Author:

Dr Michael Allen
NIHR CLAHRC South West Peninsula
University of Exeter Medical School
2.30 South Cloisters, St Luke's Campus
Heavitree Road
EXETER
EX1 2LU
Tel: 01392 726080 (mobile: 17809 297290)
E-mail: m.allen@exeter.ac.uk

Keywords: thrombectomy; stroke; health services research; health service planning

Declarations

Conflicts of interest: GAF has received personal remuneration for educational and consultancy work for Bayer, Cerevast, Medtronic, and Pfizer. GAFs institution has received a grant from Medtronic.

Funding: This study was jointly funded by the Stroke Association and the National Institute of Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care for the South West Peninsula. The views and opinions expressed in this paper are those of the authors, and not necessarily those of the Stroke Association, the NHS, the National Institute for Health Research, or the Department of Health.

Data sharing: Admissions per Lower Super Output Area (LSOA), travel times from all LSOAs to all acute stroke units, and base code used for this model may be found at https://github.com/MichaelAllen1966/stroke_unit_location

Ethical approval: Not Applicable

Guarantor: KS

Contributorship: MA, KP, MJ and KS conceived the study. MA & KP wrote the algorithm code used in this study and performed the primary analysis. MJ, PW, GAF, AGR and PM informed the base assumptions of the model, provided key literature, reviewed iterative outputs of analysis, and advised on further optimisation/analysis. MA produced the first draft of the paper, and all authoys contributed significant editing to the paper. KS oversaw all stages of the project.

Acknowledgements: We thank the Stroke Association and the NIHR CLAHRC South West for their generous support of this work.

Abstract

Purpose

Both intravenous thrombolysis (IVT) and intra-arterial endovascular thrombectomy (ET) improve the outcome of patients with acute ischaemic stroke, with ET being an option for those patients with large vessel occlusions. We sought to understand how organisation of services affects time to treatment for both IVT and ET.

Method

A multi-objective optimisation approach was used to explore the relationship between the number of IVT and ET centres and times to treatment. The analysis is based on 238,887 emergency stroke admissions in England over three years (2013-2015).

Results

Providing hyperacute care only in comprehensive stroke centres (CSC, providing both IVT and ET, and performing >150 ET per year, maximum 40 centres) in England would lead to 15% of patients being more than 45 minutes away from care, and would create centres with up to 4,300 stroke admissions/year. Mixing hyperacute stroke units (HASUs, providing IVT only) with CSCs speeds time to IVT and mitigates admission numbers to CSCs, but at the expense of increasing time to ET. With 24 CSC and all remaining current acute stroke units as HASUs, redirecting patients directly to attend a CSC by accepting a small delay (15-minute maximum) in IVT reduces time to ET: 25% of all patients would be redirected from HASUs to a CSC, with an average delay in IVT of 8 minutes, and an average improvement in time to ET of 80 minutes. The balance of CSC:HASU admissions would change from 24:76 to 49:51.

Conclusion

Planning of hyperacute stroke services is best achieved when considering all forms of acute care and ambulance protocol together. Times to treatment need to be considered alongside manageable and sustainable admission numbers.

Introduction

In England, Wales and Northern Ireland 85,000 people are hospitalised with stroke each year(1). Standard care for eligible patients with acute ischaemic stroke within 4.5 hours of onset has been intravenous thrombolysis with alteplase (IVT)(2). Time from onset to treatment is known to be especially critical, with the effectiveness of IVT declining rapidly in the first few hours after stroke(3). More recently, endovascular thrombectomy (ET) has shown substantially improved clinical outcomes in patients with large vessel occlusion (LVO) present in approximately 40% of patients with acute ischaemic stroke(4,5). ET may be effective up to 6 hours or more after stroke onset (depending on patient selection), but also demonstrates reducing effect size with increasing time from stroke onset(6). The proportion of patients eligible for ET in the UK has been estimated at about 10%(7). In England the proportion of patients receiving IVT is about 11%(1), although higher rates have been achieved as a result of reconfiguration of urban hyperacute services(1,8), with maximum rates of about 20%(1).

Providing ET presents a significant challenge for health services. The procedure is typically carried out by a neuro-interventionist and requires support from a whole theatre team (anaesthesia, nurses, radiographers) and additional imaging work up, usually with computed tomography angiography (CTA), sometimes in association with other advanced imaging techniques (CT perfusion or MRI). These additional staffing/infrastructure demands for ET require a more centralised model of service provision than currently employed for IVT(7,9).

Two models of hyperacute stroke care have been described and compared(10,11). In a mothership model all services are provided by large centres (comprehensive stroke centres, CSCs) providing both IVT and ET, at the expense of greater travel times for some patients. In a drip and ship model local IVT may be delivered at hyperacute stroke units (HASUs) before transfer of an eligible subset of patients with LVO to a CSC (ET-capable) centre. The choice between these models can depend on geography and travel times, availability of experienced staff, urban/rural split, and other factors, including the maximum practical size of a CSC under a mothership configuration, and the minimum required size of a HASU in a drip and ship model. Exploration of the advantages and disadvantages of each model may be explored by computer modelling.

Pareto-based genetic algorithms, which develop a population of solutions with varying trade-off between competing objectives have previously been established as efficient and suitable methods for addressing the complex problem of capacity-limited facility location optimisation(12). We have previously used these types of genetic algorithms to identify national configurations of hyperacute stroke services(13) that meet national guidelines recommending a minimum number of admissions to a HASU of 600 patients per year(14), coupled to the recommendation that travel time to hyperacute care should be ideally 30 minutes or less, and no more than 60 minutes(15). Here we apply that method to understanding how best ET can be provided at a national level in England, and explore how ET and IVT provision interacts. Multisociety consensus standards for thrombectomy centres have been published(16). Guidance on the minimum number of cerebral vascular procedures (40), including thrombectomy, to maintain individual operator skills has been published nationally(17), though a minimum number of thrombectomy procedures per centre in England has not yet been described. Rinaldo et al.(18) have, however, demonstrated that outcomes in high volume centres in the US, with at least 132 ET procedures per year, were better. Additionally, to provide a robust 24/7 ET service realistically requires at least 5 operators and, even with some “double scrubbing” on daytime cases, all 5 could not hope to meet minimum activity levels to maintain competence if centre volume was <150. There is no guideline on the maximum size of an

acute stroke unit, but NHS England reconfiguration guidance recommends a maximum of 1.500 admissions for single team(15), and the largest centre currently in the UK has about 2,000 admissions(1).

Method

We used a genetic algorithm based on NSGA-II(19) to derive potential configurations of stroke centres across England, balancing competing objectives. Genetic algorithms maintain a population of solutions. These solutions go through a series of generations where new solutions are formed by hybridising two existing solutions (with occasional random mutation). In each generation the best solutions are kept. We used a pareto-based method whereby, when there are multiple objectives, generated solutions are eliminated if another solution is equally as good in all optimisation parameters and is better in at least one parameter. The selected configurations were based on a range of optimisation parameters which seek to minimise travel times and to control admission numbers. Further details are given in online Appendix 1.

This location optimisation algorithm, with links to underlying data, has previously been published(13). The model predicts, for any configuration of stroke centres, the values for each of the competing objectives: travel times (estimated fastest road travel time, from home location of patient to hospital); number of admissions to each stroke centre.

We included 238,887 patients coded with ischaemic or haemorrhagic stroke (ICD-10 I61, I63, I64) with an emergency admission over a three-year period (2013-2015). Stroke admission numbers were counts of admissions for each of 31,771 Lower Super Output Areas (LSOAs) in England. No individual patient level data was accessed: counts of admissions per LSOA were extracted from Hospital Episode Statistics (HES; http://www.hscic.gov.uk/hes) with access to national HES data managed through Lightfoot Solutions (http://www.lightfootsolutions.com/). Estimated fastest road travel times were obtained from a geographic information system (Maptitude, with MP-MileCharter add-in). These travel times were used for travel time from the patient's home location to the closest stroke unit, and for travel time for onward transfer from a local HASU to the nearest CSC centre if required.

We set a preferred size for any centre providing ET as at least 150 ET procedures per year (as outcomes have been shown to be better in centres with at least 132 admissions per year(18)). In a mothership model this would equate to 1,500 admissions/year of confirmed stroke patients if 10% were eligible for ET(7). In a drip and ship model where patients may first attend a local HASU providing IVT only before onward transfer to an ET centre if appropriate, we used a minimum centre size of 600(2). As the largest stroke centre currently admits about 2,000 patients(1), we looked for solutions with the least impact on this maximum.

Outside limited randomised controlled trial data(20) there is no published evidence for what the organisational delay will be for ET in the UK, but significant transfer-related delays in ET have been described. In a large international multi-centre trial (1,000 patients across 55 sites), evaluating the use of Medtronic market-released ET devices, patients receiving ET after transfer received ET 110 minutes later than patients admitted directly, 35 minutes of which was attributable to inter-hospital travel time, suggesting a net delay of 75 minutes + travel time(21). In our modelling we have assumed some improvement over these results and have assumed a 60 minutes net delay in addition to the inter-hospital transfer travel time.

We also sought to model the impact of a preference for ambulance personnel to convey a patient with suspected acute stroke directly to an ET-capable CSC, even if a HASU (not providing ET) were closer. Again, there is no published or real-world data on the extent of this effect, but expert consensus considered that this effect may be at least 15 minutes, and we have termed this parameter 'allowable delay' i.e. an acceptable delay for all patients in arrival at hospital for the sake of a proportion who will receive ET sooner through being taken directly to a CSC. We have not modelled any attempt at the pre-hospital selection of patients with potential LVO for selective transfer directly to a CSC, but this paradigm is currently under research(22).

Results

We examined the feasibility of a mothership model by analysing the impact of having CSCs at one national centre, up to all current 127 English HASUs(1) (figure 1). With all 127 centres as CSCs the predicted average patient travel time is 18 minutes, with 90%, 98% and >99% of patients within 30, 45 and 60 minutes of their closest CSC. In this configuration, 38% of patients attend units with fewer than 600 admissions/year, and 54% of units have fewer than 600 admissions per year. If the minimum number of confirmed stroke admissions required to sustain an ET service is 1,500 (resulting in ~150 ET procedures per year) then the maximum number of CSCs from the solutions identified by the algorithm is 40. The fastest average travel time for solutions with at least 1,500 admissions per year is 29 minutes, with 62%, 85% and 95% within 30, 45 and 60 minutes of their closest centre. Under these parameters, the largest of the 40 CSCs would receive about 3,000 confirmed stroke admissions per year.

We examined the distribution of unit sizes in the obtained solutions. For example, 69 solutions were identified with 30 CSCs that also had a minimum number of admissions of 1,500. The range in admission numbers per year across all solutions was 1,515 to 5,722 (interquartile range [IQR] 2,071-3,246).

The maximum number of CSCs is highly sensitive to the proportion of patients eligible for ET and the minimum recommended thrombectomy procedures. In our base case we have assumed that a CSC must have a catchment of at least 1,500 stroke admissions per (derived from an estimate of 10% receiving thrombectomy, and 150 procedures per year being the acceptable minimum). This limits the number of CSCs to 40 (assuming there are no other constraints). These patients may all attend the CSC directly, or those requiring thrombectomy may arrive via a HASU. At this point 62% of patients are within 30 minutes of a CSC. Figure 1 (middle panel) shows the how the maximum number of CSCs would change if the required catchment changed. The required catchment may change because of differences in thrombectomy rate, or changes to the minimum acceptable number of thrombectomy procedures carried out. For example, if a CSC only required a catchment of 1,000 patients per year there could be 57 CSCs (if no other constraints exist) with 72% of patients within 30 minutes. However, if a a CSC required a catchment of 2,000 patients per year, the maximum number of CSCs would be 30, with 57% patients being within 30 minutes.

The most practical starting point for the location of ET centres is the 24 current English neurosciences centres, which already have neurointerventional staff and the necessary imaging and interventional suite infrastructure to deliver ET. A mothership model based solely on these centres is predicted to have an average travel time of 38 minutes, with 43%, 71% and 86% of patients within 30, 45 and 60 minutes of their closest centre. The number of admissions to each centre would range from 1,264 to 6,117 (IQR: 2,292-4,169). We compared these performance statistics with other configurations of 24 centres (if those centres could be chosen from any of the current 127

English HASUs). We made no judgement in the model on the practicality of "relocating" any [neuroscience] centre. The fastest average travel times for any 24-unit configuration would be of 33 minutes, with 48%, 80% and 93% of patients within 30, 45 and 60 minutes of their closest centre, with admissions ranging from 1,595 to 7,639 (IQR: 2,092-3,849). Admissions could be more even distributed by choosing an alternative 24-centre scenario which produces an average travel time of 36 minutes, with 48%, 73% and 88% of patients within 30, 45 and 60 minutes of their closest centre, and with admissions ranging from 2,236 to 4,382 (IQR: 2,732-3,978). However, in all these scenarios the upper limit of the pre-set desirable admission range (600-2,000) is substantially exceeded in many centres.

When looking for feasible mothership solutions we have assumed an upper limit of size of 2,000 admissions per year. If this is the case, then at least 55 units would be needed across England which cannot be reconciled with the maximum number of 40 units required to sustain 1,500 admissions per year to all CSCs. If the minimum unit size must be 1,500 admissions per year (and no HASUs are used) then solutions exist with largest unit admissions per year in the range 2,750 to 3,000.

If a pure mothership model is considered unfeasible, then HASUs need to co-exist with the CSCs within a drip and ship model of care, whereby patients thought likely to be suitable for ET would receive IVT at a HASU (if that centre is their closest centre) and would then be transferred to the nearest CSC. In order to explore the impact of adding HASUs to a configuration of CSCs (moving away from a pure mothership model towards a drip and ship model) we took the current 24 neuroscience centres in England as a baseline CSC configuration, and sequentially added HASUs centres from the subset of 103 remaining HASU locations. We assumed, initially, that patients travel first to their closest centre of any type. All patients eligible for IVT would receive IVT at this first centre, but those with LVO likely to benefit from ET will then be transferred to the closest ET capable CSC. In this configuration, we assumed a net delay to ET of 60 minutes (excluding travel time from the CSC to the HASU) for those patients taken first to a HASU.

If additional HASUs are chosen to minimise the average time to arrival at the first centre, increasing the number of HASUs reduces average time to IVT, but increases the average time to ET (figure 2). With only 24 CSCs in a mothership model, average time to arrival for ET and IVT is 38 minutes, with 71% of patients arriving within 45 minutes travel time. If at the other extreme, the remaining 103 current English units are present as HASUs in a drip and ship model, the average time to arrival for IVT is reduced from 38 to 18 minutes, but the average time to arrival for ET is increased from 38 to 96 minutes, with only 24% of patients arriving within 45 minutes of stroke onset. The complex relationship between the number of HASUs, travel times and admission numbers for the first hospital attended is shown in figure 3.

If all acute stroke centres have a minimum of 600 admissions/year then the maximum number of additional HASUs would be 58 (82 centres in total) and all first hospital admissions would be between 600 and 1,810 per year (IQR: 781-1,119). Average time to arrival for IVT falls from 38 minutes (with 24 neuroscience centres only) to 22 minutes, with 80%, 94% and 98% of patients within 30, 45 and 60 minutes. However, the average time to ET would increase from 38 to 89 minutes, with 31% of patients arriving within 45 minutes travel time (travel times include an additional 60 minutes net organisational delay). The algorithm identified 988 solutions where annual admissions to all centres are within the range 600 to 2,000; these solutions have between 57 and 82 centres in total (from which the 24 existing neurosciences centres provide ET). We examined the distribution of unit sizes in these obtained solutions. The range in admission numbers per year across all solutions was 601 to 2,000 (IQR: 911-1,394).

We anticipated that in a mixed system of HASU and CSCs, there may be a preference to convey suspected stroke cases directly to the CSC, even if a HASU were closer, a phenomenon we have called ‘allowable delay’. A 15-minute allowable delay would mean that the ambulance would transfer a patient directly to a more distant CSC so long as there was no more than 15 minutes extra travel time.

We modelled the impact of ‘allowable delay’ in the drip and ship model, with 24 CSCs located at the current neuroscience centres and with 103 HASUs at all the remaining centres. As before, assessment and IVT at a nearer HASU incurred a net delay of 60 minutes plus transport time. In the model, patients could be taken straight to a CSC with an allowable delay in IVT of 0-100 minutes, applied to all patients. Generally, allowing a delay in IVT to directly attend a CSC increases average time to IVT and decreases average time to ET (figure 4). However, the effects are not equal; allowing up to 15 minutes delay for IVT increases average time to IVT by just 2 minutes while reducing average time to ET by 20 minutes. This allowable delay affects 25% of all patients, who have an average delay in IVT of 8 minutes, and an average improvement in time to ET of 80 minutes.

The practice of accepting a delay in IVT for the sake of directly attending a CSC was found to have a significant effect on admission numbers to hospitals (figure 5). As the allowable delay in IVT increases, admissions to CSCs increase while admission numbers to HASUs reduce. With a 15-minute allowable delay for IVT, the number of centres with fewer than 300 admissions per year increases from 7 to 30, while the number of centres with fewer than 600 admissions per year increases from 69 to 92 (out of 127 HASUs). At the same time the number of centres with more than 2,500 admissions increases from none to five. With a 30-minute allowable delay for IVT the number of centres with fewer than 300 IVT admissions per year increases further to 68, the number with fewer than 600 IVT admissions per year increases to 99, and the number with more than 2,500 IVT admissions increases to 11. The balance of CSC:HASU admissions would change from 24:76 to 49:51.

The effect of mixing HASUs with 24 CSCs was examined in more detail with an allowable delay in IVT of 15 minutes to directly attend a CSC (figure 6). The maximum total number of centres that can maintain at least 600 admissions per year is 54 (reduced from 82 in configurations with no allowable delay in IVT). With this configuration the average time to IVT is 26 minutes with 67%, 91% and 97% of patients within 30, 45 and 60 minutes of the first admitting centre. The average time to ET is 78 minutes (61% of patients within 45 minutes and 65% within 60 minutes). Admission numbers to the first admitting hospital range from 610 to 4,936 (IQR: 789-1,840).

The on-line appendix contains example maps and more detailed analysis of some specified configurations: a 24 CSC mothership model, a 30 CSC mothership model, and a 30 CSC + 50 HASU drip and ship model with either 0 or 15-minute allowable delay in IVT.

Discussion

In an ideal configuration of hyperacute stroke care, all patients would live close to a CSC offering high quality acute stroke unit care and both IVT and ET. Our results suggest that in England, providing all acute stroke care in such centres is not likely to be considered feasible. There are three limitations to the number of CSCs: 1) economic, 2) time to train required staff in thrombectomy procedures, and 3) admission numbers (beyond a certain number of CSCs some units will fall below the admission numbers anticipated to be required to maintain a 24/7 expert thrombectomy service). With our base assumptions (10% of patients eligible for ET, and minimum

CSC size delivering >150 ET/year) the algorithm identified that configurations would have up to 40 centres, however, 15% of stroke patients (over 12,000 patients annually) would be further than 45 minutes away from such centres, and centres would have to cope with 3,000 or more confirmed stroke admissions per year (the largest stroke centre in England currently admits just over 2,000 patients per year(1)). Additionally, when considering substantial service reconfiguration there is some uncertainty over the exact proportion of stroke patients who will receive ET.

On this basis, a mixed model of CSCs (offering both IVT and ET) and HASUs (offering only IVT) appears necessary. Such a model of care reduces time to admission at the first centre (providing IVT) but will delay ET for many eligible patients. The delay comes from additional transport time and from organisational delays in arranging/starting onward travel(21). The organisational delay could be reduced by having ambulances wait at the first-admission hospitals to see if CTA indicates that ET is required, or by prioritising ambulance provision for transfer to a CSC. Consideration of the substantial impact on ambulance services would need to be given with either of these strategies(23).

The balance between HASUs and CSCs will vary across the country. Where population is dense (such as London and other large cities), it is likely that services may delivered entirely through CSCs. Where population is less dense, in areas with more rural populations, more HASUs will be required to control the time to admission to first hospital.

When planning both IVT and ET services, time to treatment for both procedures and broader issues of access to sustainable high quality hyperacute stroke services must be considered. It is possible that there could be an ‘over-supply’ of more local IVT services where time to ET is increased significantly with minimal improvements in time to IVT. However local HASUs may be required, if only to mitigate the number of direct admissions to CSCs.

Our modelling suggests that allowing a small delay in IVT for patients to be taken straight to a CSC may significantly reduce time to ET without significant effect on time to IVT. This is similar to an analysis by Froehler *et al*. who found, in a hypothetical bypass analysis, if patients were brought directly to the ET-capable centre, IVT would be slightly delayed (by 12 minutes) but ET would be delivered 91 minutes sooner(21). However, the population sizes are different – about 30% of all patients would have some delay in initial assessment and treatment (with IVT for up to 1 in 5 of these patients), whereas only about 1 in 10 of these diverted patients would be expected to benefit from improved time to ET. Ideally, we would like to model expected clinical outcomes as times to thrombolysis and thrombectomy are changed. At the moment, however, there is not the available data or analysis to model clinical outcome as a function of both time to thrombolysis and time to thrombectomy.

Our modelling also anticipates the practicality of implementing a mixed model of HASUs and CSCs, in that there may be a preference to convey a patient with suspected stroke to a CSC that is more distant, so long as the additional travel time is not excessive. We have shown that such an ‘allowable delay’ may significantly reduce time to ET without a substantial adverse effect on average time to IVT. Though models of clinical outcome are in relatively early stages of development, a probabilistic model of good outcome is being developed whereby for any individual patient it may be determined whether a good outcome is more likely to be achieved by going straight to an ET-capable centre rather than first attending a local IVT-only centre(10,11). There is, however, a potential conflict between making the best decision for any individual patient, and making decisions that destabilise the wider services for all patients. Holodinsky et al.(11) examined the likely outcome of drip and ship vs. mothership models of care, focussing on decisions which

maximise the likelihood of a good clinical outcome for any individual patient. Such decisions, however, may undermine the sustainability of stroke services, either by undermining the capability of a local HASU by falling below minimum recommended levels of activity, or by overwhelming the capacity of CSCs with very large numbers of suspected stroke patients. The reduction in the size of centres may be mitigated, at least in part, by additional expert support to HASUs provided by the CSCs, such as the use of telemedicine/teleradiology to support clinical decision making(24). By the same token, the large numbers of admissions directly to a CSC may be mitigated by the pre-hospital selection of patients more likely to have a LVO, and the use of such instruments is currently under investigation(22).

We have modelled on the basis that determination of suitability for thrombectomy can only be made at a hospital. There are however attempts to improve pre-hospital decision-making, ranging from pre-hospital triage scales(22), through biochemical tests(25), through to ambulances equipped with imaging equipment(26). It is beyond the scope of this paper to deal with the effects of such pre-hospital systems, but this is an area where further modelling should be of value.

It is not only provision of ET that may pull in more patients to a CSC. Advanced imaging techniques available there 24/7, but not necessarily elsewhere, may identify patients who might benefit from ET whose stroke onset time is either unknown or is longer than would be considered for IVT(27,28).

Our model makes assumptions about, and simplifications of, the real world. Indeed, modelling may be thought of as the art of abstracting a messy real-world application into a clean problem suitable for an algorithmic solution. While not detracting from the key learnings, these simplifications need to be born in mind when interpreting our findings. A key assumption is that patients will be taken to the closest appropriate stroke centre to their home. We have previously shown that our model generally has very good predictability for admission numbers, but the prediction of admission numbers is poorer when centres are located close to each other(13). When comparing predicted with actual admissions there was a median absolute error of 105 admissions per centre per year, or a relative absolute error of 17%, but the error was typically about 12% when centres were separated by 30 minutes or more travel time, and 20-30% when centres were separated by less than 30 minutes. This inaccuracy may reflect considerations other than expected travel time, such as on-the-day traffic conditions, or may reflect a person having a stroke close to, but not actually, at home, such as a place of work. When centres are more closely located such as in city metropolitan areas like London, there may therefore be more potential to smooth admissions across centres without a significant detrimental effect on travel time.

We have assumed that once at the destination hospital, all hospitals will perform equivalently in speed to treatment. Bray et al.(29) have noted, however, that larger hospitals (those performing at least 50 thrombolysis procedures per year) delivered treatment 20-30 minutes faster than smaller hospitals (performing <50 thrombolysis procedures per year). In the objectives for our modelling we have sought to avoid small units, but it is still possible that larger units will be better able to deliver treatment faster or more consistently across the day and week. Decisions on whether a patient is a suitable candidate for thrombectomy depend on the patient having a CT-angiogram. While it is anticipated that all HASUs in England will be able to perform this, having rapid time to analyse and report on scans may be more challenging for smaller units (though this may be mitigated, at least in part, by alternative strategies, such as telemedicine).

We have not taken into account the admission of stroke mimics to centres in these models. Recent experience evidence suggests that about 25% of admissions to an acute stroke unit are subsequently

identified as a stroke mimic(30). In large volume centres, stroke mimics have significant impact on bed utilisation and infrastructure support. Improvements in pre-hospital diagnosis of ischaemic stroke may in the future lessen the potential impact of stroke mimics on centralised stroke services.

We have also not considered ischaemic stroke patients who are not recognised to have stroke prior to hospital admission. The majority, but not all, ischaemic stroke patients eligible for IVT and nearly all patients suitable for ET are FAST positive(31). However, these patients may be admitted to hospitals without HASUs and require secondary transfer to hospitals with a HASU.

We have not taken growth in stroke incidence into account in this analysis. With an ageing population, we anticipate a significant increase in admissions to hospital with disabling stroke despite better preventative care, particularly in stroke related to atrial fibrillation[24]. Although such forecasting is imprecise, a potential increase in stroke incidence and hospital admissions could be driven by a predicted 54% increase in the population of England aged 75 or over the next 15 years(32). These increases may have significant implications for centres close to the lower margin of institutional activity in current recommendations.

Though models always have simplifications, we believe this analytic approach provides clear guidance regarding provision of ET services. In England, a mothership model based on providing hyperacute stroke care only in CSCs does not look feasible, and so consideration must be given to the optimal configuration of a drip and ship model with local HASUs providing IVT and transferring eligible patients to a tertiary CSC. In such a mixed model the HASUs not only help to provide more rapid IVT, but they also mitigate admission numbers that directly attend CSCs, thus preventing overload of the tertiary CSC and helping maintain flow through patient pathways. However, an over-reliance on local HASUs may significantly increase time to ET with little benefit to time to IVT, and may also exceed the optimum number of HASUs necessary to control direct admissions to CSCs. When considering individual patients it may appear beneficial to accept a short delay in IVT for the sake of direct admission to a more distant CSC for consideration of ET. However when considering the overall population disability benefit, this strategy should be applied carefully as it has the potential to destabilise both IVT and ET provision by distorting admission numbers to both types of centre, risking making many HASUs too small to be sustainable and CSC admissions too large to be manageable, to the detriment of all stroke patients not just those eligible for IVT &/or ET. Improvements in pre-hospital diagnosis of stroke due to LVO and stroke mimics would enable such a strategy to be implemented with less destabilising effect.

Conclusion

Planning of hyperacute stroke services is best achieved when considering all forms of acute care, and ambulance protocol, together. Times to treatment need to be considered alongside manageable and sustainable admission numbers to different types of acute stroke centre.

References

1. Sentinel Stroke Audit Programme (SSNAP). Apr2015Mar2016-AnnualResultsPortfolio. National Results April 2015-Mar 2016 https://www.strokeaudit.org/Documents/Results/National/Apr2015Mar2016/Apr2015Mar2016-AnnualResultsPortfolio.aspx. 2016.
2. Intercollegiate Stroke Working Party. National Clinical Guideline for Stroke, 5th Edition (Royal College of Physicians of London). [Internet]. 2016 [cited 2017 Aug 2]. Available from: www.strokeaudit.org.uk/guideline
3. Emberson J, Lees KR, Lyden P, Blackwell L, Albers G, Bluhmki E, et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: A meta-analysis of individual patient data from randomised trials. Lancet. 2014;384(9958):1929–35.
4. Flynn D, Francis R, Halvorsrud K, Gonzalo-Almorox E, Craig D, Robalino S, et al. Intra-arterial mechanical thrombectomy stent retrievers and aspiration devices in the treatment of acute ischaemic stroke: A systematic review and meta-analysis with trial sequential analysis. Eur Stroke J [Internet]. 2017;2(4):308–18. Available from: http://journals.sagepub.com/doi/10.1177/2396987317719362
5. Smith WS, Lev MH, English JD, Camargo EC, Chou M, Johnston SC, et al. Significance of large vessel intracranial occlusion causing acute ischemic stroke and tia. Stroke. 2009;40(12):3834–40.
6. Goyal M, Menon BK, Van Zwam WH, Dippel DWJ, Mitchell PJ, Demchuk AM, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: A meta-analysis of individual patient data from five randomised trials. Lancet. 2016;387(10029):1723–31.
7. McMeekin P, White P, James MA, Price CI, Flynn D, Ford GA. Estimating the number of UK stroke patients eligible for endovascular thrombectomy Estimating the number of UK stroke patients eligible for endovascular thrombectomy. Eur Stroke J. 2017;2:319–26.
8. Morris S, Hunter RM, Ramsay AIG, Boaden R, McKevitt C, Perry C, et al. Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis. Bmj [Internet]. 2014;349(aug04 4):g4757–g4757. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.g4757
9. Muir KW, White P. HERMES: Messenger for stroke interventional treatment. Vol. 387, The Lancet. 2016. p. 1695–7.
10. Holodinsky JK, Williamson TS, Kamal N, Mayank D, Hill MD, Goyal M. Drip and ship versus direct to comprehensive stroke center. Stroke. 2017;48(1):233–8.
11. Holodinsky JK, Williamson TS, Kamal N, Mayank D, Hill MD, Goyal M. Drip and ship vs. direct to comprehensive stroke centre: the effect of large vessel occlusion screening tools on conditional probability modeling. Int J Stroke. 2017;12(4_suppl):4–84 (ACU.11).

12. Gauss MP, Sarajevo C. Application of the Pareto-Based Genetic Algorithm in Multi-Objective Location Analysis Application of the Pareto-Based Genetic Algorithm in Multi- Objective Location Analysis. 33rd Int Conv MIPRO. 2010;680–5.

13. Allen M, Pearn K, Villeneuve E, Monks T, Stein K, James M. Feasibility of a hyper-acute stroke unit model of care across England. A modelling analysis. BMJ Open [Internet]. 2017;doi.org/10.1136/bmjopen-2017-018143. Available from: http://bmjopen.bmj.com/cgi/content/full/bmjopen-2017-018143?ijkey=uI4cCifsPKLccU0&keytype=ref

14. Price C, James M. Meeting the Future Challenge of Stroke Meeting the Future Challenge of Stroke [Internet]. British Association of Stroke Physicians. 2011. Available from: https://basp.ac.uk/wp-content/uploads/2017/02/BASP-Meeting-the-Future-Challenge-of-Stroke-2011-15.pdf

15. NHS England. Stroke Services: Configuration Decision Support Guide [Internet]. 2015. Available from: http://www.necn.nhs.uk/wp-content/uploads/2015/02/ADD_1210_Stroke-NHS-toolkit_c1_full_2016.08.18_11.48_10_single-1.pdf

16. White PM, Bhalla A, Dinsmore J, James M, McConachie N, Roffe C, et al. Standards for providing safe acute ischaemic stroke thrombectomy services (September 2015). Clin Radiol. 2017;72(2):175.e1-175.e9.

17. Lenthall R, Mcconachie N, White PP, Clifton A, Group UKN, Society B. BSNR training guidance for mechanical thrombectomy. Clin Radiol [Internet]. Elsevier; 2017;72(2):175.e11-175.e18. Available from: http://dx.doi.org/10.1016/j.crad.2016.11.007

18. Rinaldo L, Brinjikji W, Rabinstein AA. Transfer to High-Volume Centers Associated with Reduced Mortality after Endovascular Treatment of Acute Stroke. Stroke. 2017;48(5):1316–21.

19. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput. 2002;6(2):182–97.

20. Muir KW, Ford GA, Messow C-M, Ford I, Murray A, Clifton A, et al. Endovascular therapy for acute ischaemic stroke: the Pragmatic Ischaemic Stroke Thrombectomy Evaluation (PISTE) randomised, controlled trial. J Neurol Neurosurg Psychiatry [Internet]. 2017;88:38–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27756804

21. Froehler MT, Saver JL, Zaidat OO, Jahan R, Aziz-Sultan MA, Klucznick RP, et al. Interhospital Transfer Prior to Thrombectomy is Associated with Delayed Treatment and Worse Outcome in the STRATIS Registry. Circulation [Internet]. 2017; Available from: https://doi.org/10.1161/CIRCULATIONAHA.117.028920%0Ahttp://circ.ahajournals.org/lookup/doi/10.1161/CIRCULATIONAHA.117.028920

22. Schlemm L, Ebinger M, Nolte CH, Endres M. Impact of Prehospital Triage Scales to Detect Large Vessel Occlusion on Resource Utilization and Time to Treatment. Stroke [Internet]. 2018;49(2):439–46. Available from: http://stroke.ahajournals.org/lookup/doi/10.1161/STROKEAHA.117.019431

23. Ng FC, Low E, Andrew E, Smith K, Campbell BCV, Hand PJ, et al. Deconstruction of Interhospital Transfer Workflow in Large Vessel Occlusion. In: Stroke. 2017. p. 1976–9.

24. Wechsler LR, Demaerschalk BM, Schwamm LH, Adeoye OM, Audebert HJ, Fanale C V., et al. Telemedicine quality and outcomes in stroke: A scientific statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2017;48(1):e3–25.

25. Tian F, Bibi F, Dale N, Imray CHE. Blood purine measurements as a rapid real-time indicator of reversible brain ischaemia. Purinergic Signal. 2017;13(4):521–8.

26. Fassbender K, Grotta JC, Walter S, Grunwald IQ, Ragoschke-Schumm A, Saver JL. Mobile stroke units for prehospital thrombolysis, triage, and beyond: benefits and challenges. Vol. 16, The Lancet Neurology. 2017.

27. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. N Engl J Med [Internet]. 2017; Available from: http://www.nejm.org/doi/10.1056/NEJMoa1706442

28. Campbell BCV, Mitchell PJ, Kleinig TJ, Dewey HM, Churilov L, Yassi N, et al. Endovascular Therapy for Ischemic Stroke with Perfusion-Imaging Selection. N Engl J Med. 2015;372(11):1009–18.

29. Bray BD, Campbell J, Cloud GC, Hoffman A, Tyrrell PJ, Wolfe CDA, et al. Bigger, faster?: Associations between hospital thrombolysis volume and speed of thrombolysis administration in acute ischemic stroke. Stroke. 2013;44(11):3129–35.

30. Dawson A, Cloud GC, Pereira AC, Moynihan BJ. Stroke mimic diagnoses presenting to a hyperacute stroke unit. Clin Med. 2016;16(5):423–6.

31. Harbison J, Hossain O, Jenkinson D, Davis J, Louw SJ, Ford GA. Diagnostic accuracy of stroke referrals from primary care, emergency room physicians, and ambulance staff using the face arm speech test. Stroke. 2003;34(1):71–6.

32. Office of National Statistics. Office of National Statistics (2014). 2014-based subnational population projections. [Internet]. 2014. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/datasets/regionsinenglandtable1

Figures

Figure 1. Feasibility of a mothership model where all patients attend their closest comprehensive stroke centre which provides both IVT and ET. The panels show the relationship between the number of centres and best identified results for: (left panel) average and maximum travel time to centre, (centre panel) greatest and fewest admissions per year to any single centre, and (right panel) the proportion of patients attending a centre within 45 minutes and providing at least 150 ET procedures per year.

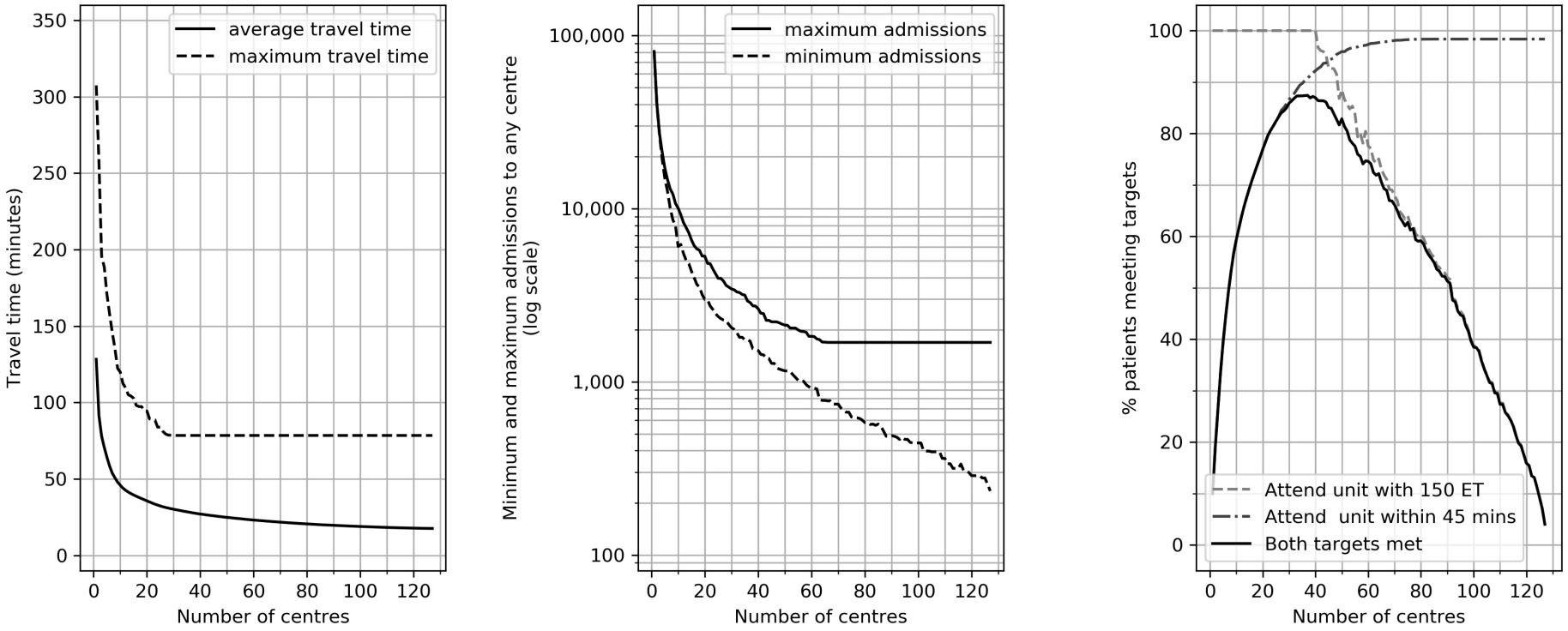


Figure 2. Travel times drip n ship model, where all patients first attend closest centre, with onwards travel to comprehensive stroke centre (CSC) if patient requires endovascular thrombectomy (ET). The base case has 24 CSCs (located in current 24 neuroscience centres), with Hyper Acute Stroke Units (HASUs) providing only intravenous thrombolysis (IVT) being added from the remaining 103 existing acute stroke centres in order to minimise average travel time to the first hospital. Showing the effect of adding HASUs (that only offer IVT) on the average time to arrival at hospital for IVT and ET. Onward travel for ET includes the transfer travel time and an additional 60-minute transfer-related delay in receiving ET.

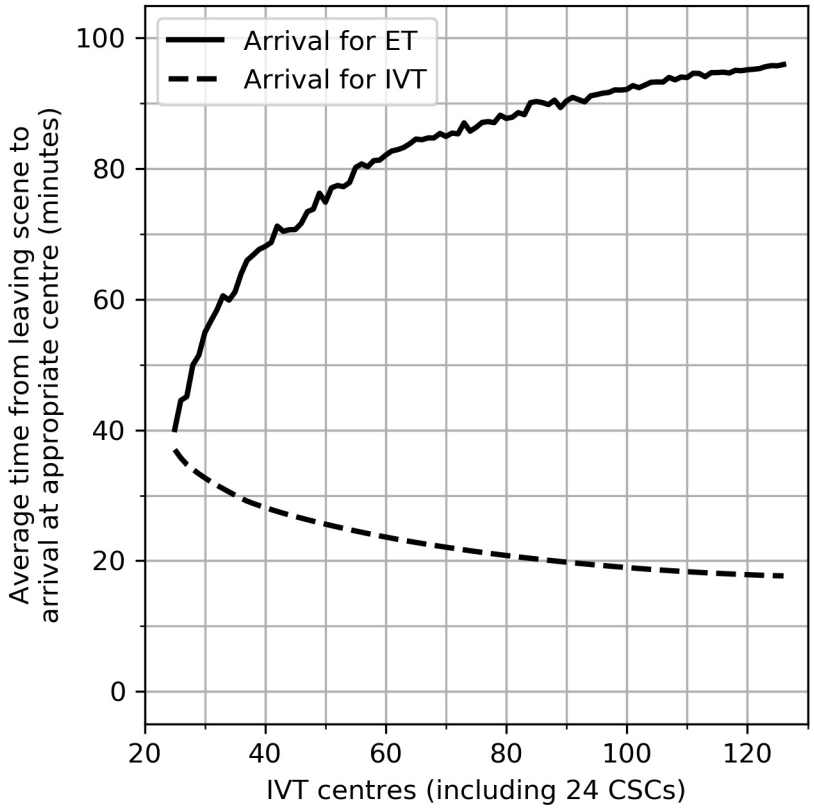


Figure 3. Feasibility of a drip n ship model, where all patients first attend closest centre, with onwards travel to comprehensive stroke centre (CSC) if patient requires endovascular thrombectomy (ET). The base case has 24 CSCs (located in current 24 neuroscience centres), with Hyper Acute Stroke Units (HASUs) providing only intravenous thrombolysis (IVT) being added from the remaining 103 existing acute stroke centres in order to minimise average travel time to the first hospital. Showing the effect of adding HASUs on the arrival characteristics of the first hospital attended (for IVT treatment). The panels show the relationship between the number of IVT centres and best identified results for: (left panel) average and maximum travel time to first hospital attended, (centre panel) greatest and fewest admissions per year to any single centre, (right panel) the proportion of patients attending a centre within 30 minutes and admitting at least 600 confirmed stroke patients per year.

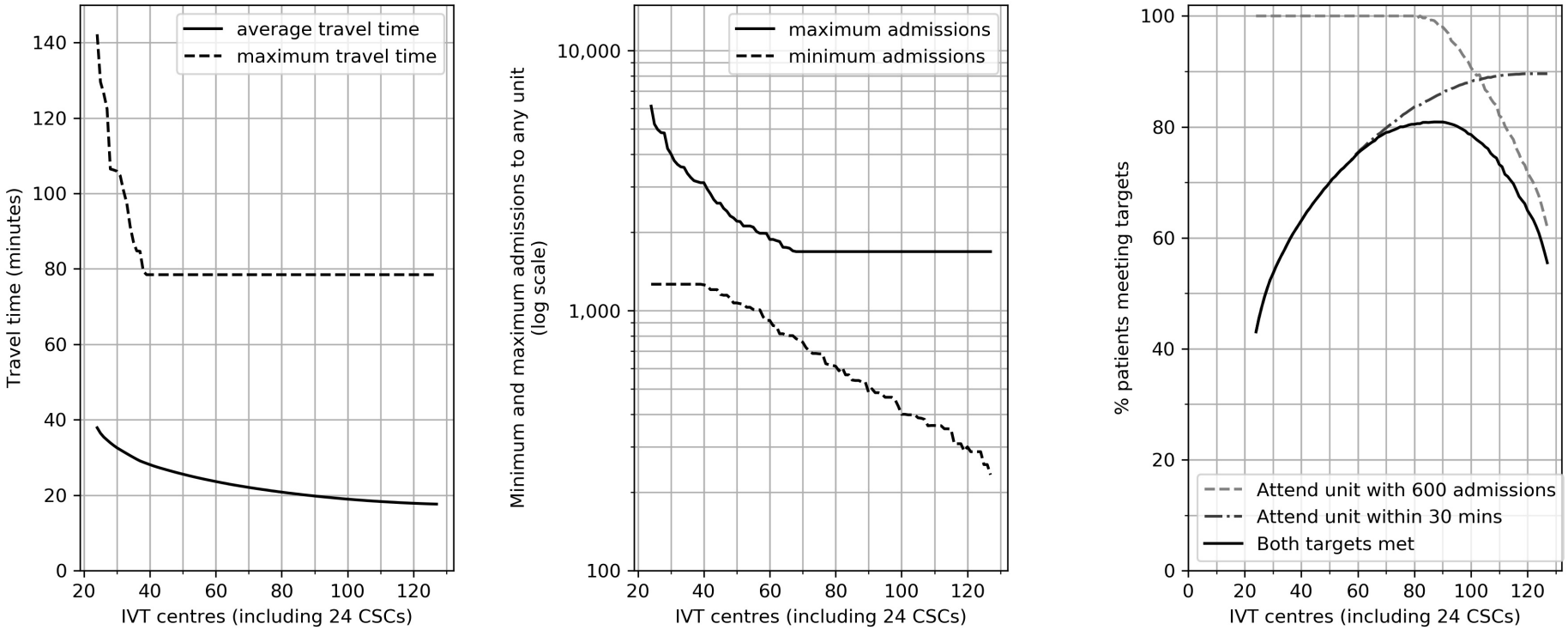


Figure 4. A drip n ship model where all patients first attend closest centre, with onwards travel to comprehensive stroke centre (CSC) if patient requires endovascular thrombectomy (ET). Taking a fixed configuration (the current neuroscience centres as the 24 CSCs and the remaining 103 existing acute centres as Hyper Acute Stroke Units, HASUs), the chart shows the effect of allowing a time delay in receiving intravenous thrombolysis (IVT) in order to travel directly to a CSC (capable of both IVT and ET, thus saving transfer delays for ET). The lines show average time to arrival for ET (solid line) and IVT (dashed line). Onward travel for ET includes the transfer travel time and an additional 60-minute transfer-related delay in receiving ET.

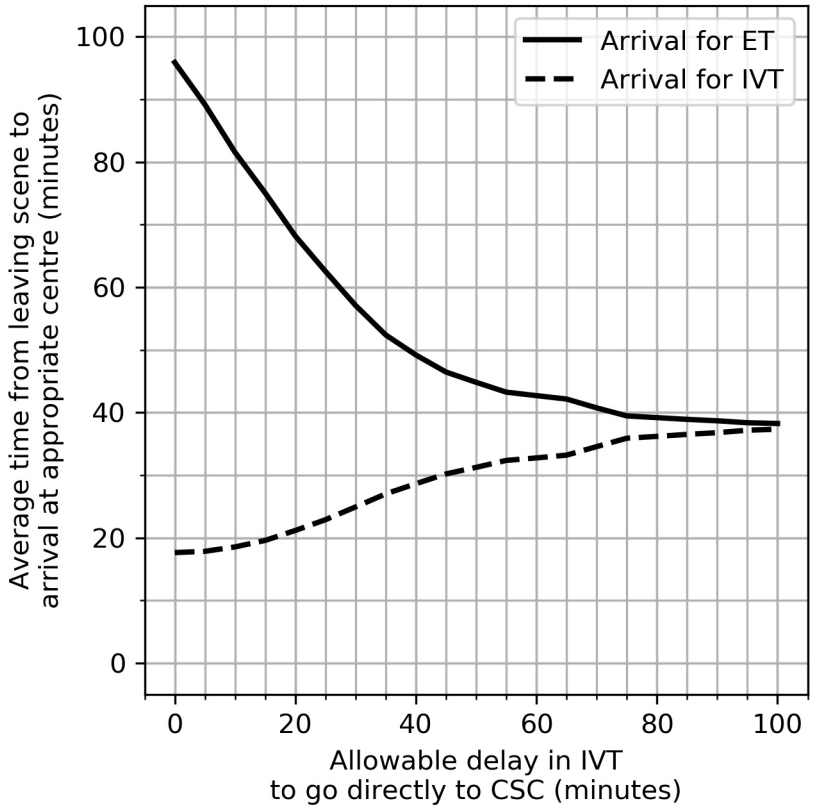


Figure 5. A drip n ship model where all patients first attend closest centre, with onwards travel to comprehensive stroke centre (CSC) if patient requires endovascular thrombectomy (ET). Taking a fixed configuration (the current neuroscience centres as the 24 CSCs and the remaining 103 existing acute centres as Hyper Acute Stroke Units, HASUs) the panels show the effect of allowing a time delay in receiving intravenous thrombolysis (IVT) in order to travel directly to a CSC (capable of both IVT and ET, thus saving transfer delays for ET). The violin plot (left panel) shows the effect on the admissions to the first admitting centres: range, median (middle bar) and distribution (shaded body). The line chart (right panel) shows the effect on the number of centres below or above a given threshold of annual admissions.

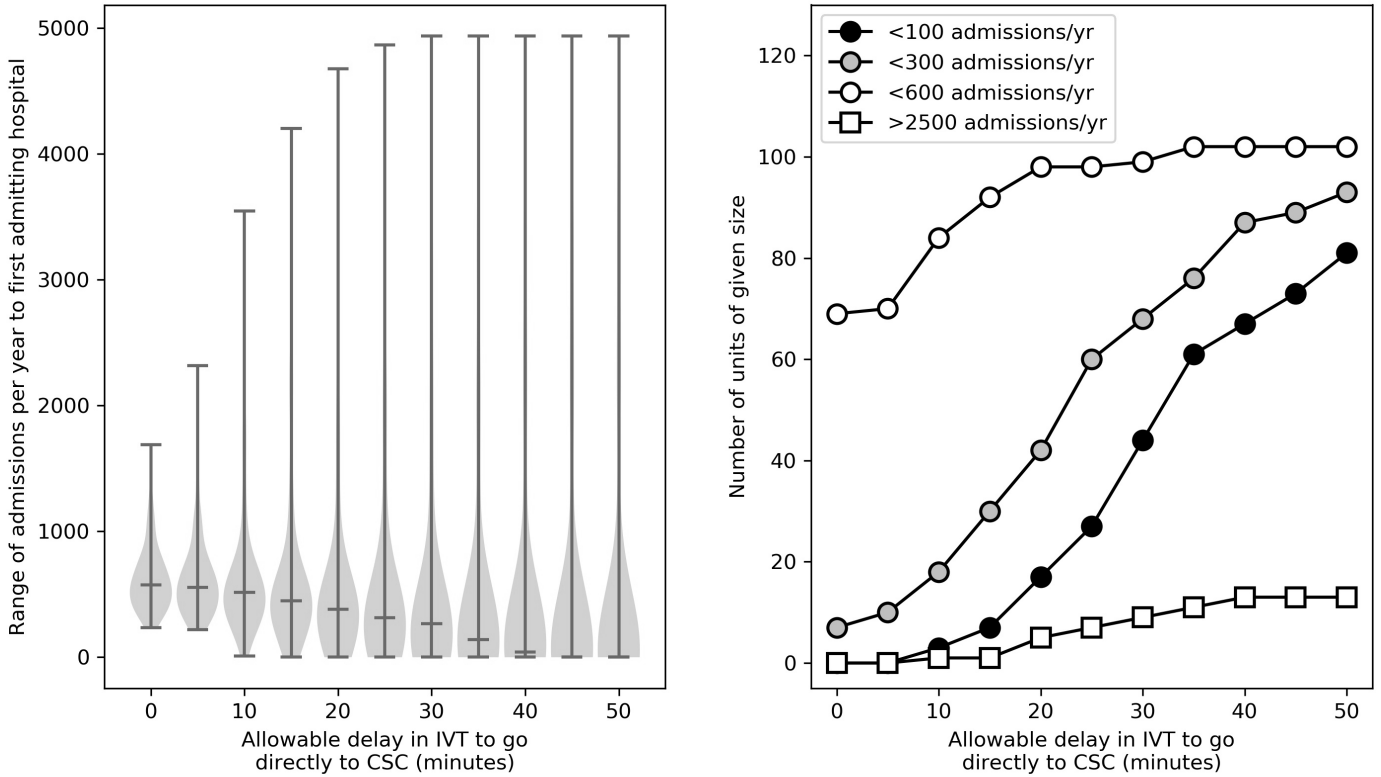
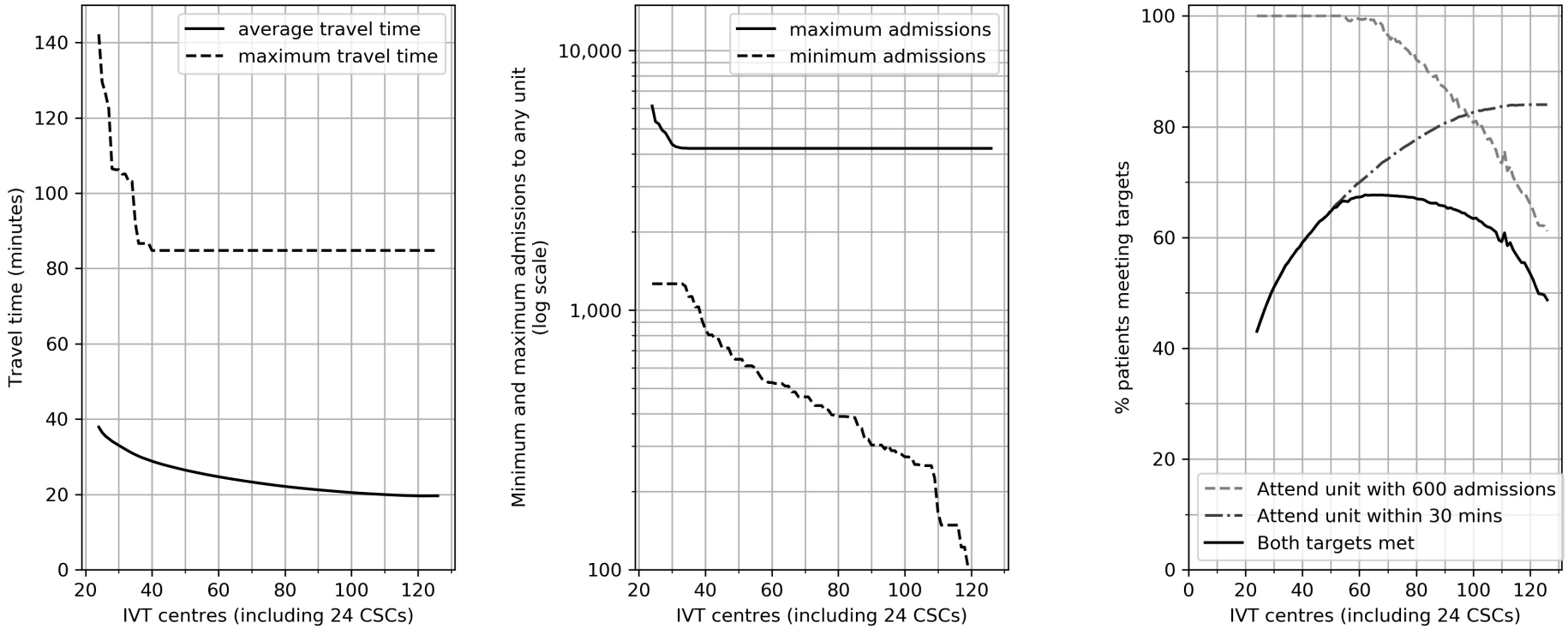


Figure 6. Feasibility of a drip n ship model, where all patients first attend closest centre, with onwards travel to comprehensive stroke centre (CSC) if patient requires endovascular thrombectomy (ET), assuming an allowable intravenous thrombolysis (IVT) delay of up to 15 minutes in order to travel directly to a CSC (capable of both IVT and ET, thus saving transfer delays for ET). The base case has 24 CSCs (located in current 24 neuroscience centres), with Hyper Acute Stroke Units (HASUs) being added from the remaining 103 existing acute stroke centres in order to minimise average travel time to the first hospital. Showing the effect of adding HASUs on the arrival characteristics of the first hospital attended (for IVT treatment). The panels show the relationship between the number of IVT centres and best identified results for: (left panel) average and maximum travel time to CSC, (centre panel) greatest and fewest admissions per year to any single CSC, (right panel) the proportion of patients attending a CSC within 30 minutes and having at least 600 confirmed stroke admissions per year.



Appendix 1: Methodology

1 Description of problem

In order to establish a hyper-acute stroke unit (HASU) model for emergency stroke care across England, all HASUs (providing thrombolysisis) should ideally have a minimum of 600 yearly admissions of confirmed strokes. Comprehenisve stroke centres (providing thrombolysis and thrombectomy) should perform a minimum of 150 thrombectomy proecures per year (10% of stroke patients are assumed to receive thrombectomy). No unit should be infeasibly large, and we have taken the current largest unit with ~2,000 stroke admissions per year as our upper target, though the algorithm also explores use of larger units.

The problem involves looking for solutions that can place any number of hospitals in any of 127 locations. There are therefore 2^127 or 10^38 possible solutions. Each solution requires looking up road travel times from each of 31,171 patient locations to all open hospitals to allocate patients to their closest hospital. There are 13 possible objectives to try to achieve or trade-off (see section 3.1).

This type of problem is termed ‘NP-hard’ - it cannot be solved explicitly in reasonable time. And as there are multiple-objectives that trade-off against each other there is no single solution to the problem (as there is no way to objectively determine the weighting of different objectives); rather we are looking for a population of solutions which demonstrate the trade-off between different objectives.

With NP-hard problems there are often a range of different heuristic algorithms which search for good solutions to the problem, while never guaranteeing an optimal solution is found. One set of general purpose heuristic methods are a family of algorithms known as ‘genetic algorithms’, due to their inspiration coming from the theory of evolution. Here we describe the specific genetic algorithm used in our study.

2 Code and data repository

Data and code used for the model are available at:
https://github.com/MichaelAllen1966/1805_thrombectomy_location

Note: The code contains a bespoke Genetic Algorithm written in Python/NumPy. No Genetic Algorithm libraries were used.

3 Multi-objective problem

3.1 Pareto dominance

When solving an optimisation problem based on one objective, the optimal solution is given by the configuration with the best (highest or lowest) objective value. In the case of multi-objective optimisation, comparing several solutions requires to reference to the notion of dominance: a vector $a$of the objective space dominates another vector b if all criteria of a are better or equal to criteria of b and $a \neq b$[1]. Then, a solution is non-dominated if there are no other solutions at least equal in all objectives and better in at least one objective. At the end of the optimisation process, there is no

single best solution but a set of non-dominated solutions, called the Pareto Front. An example of a Pareto Front using two objectives is shown in figure 1.

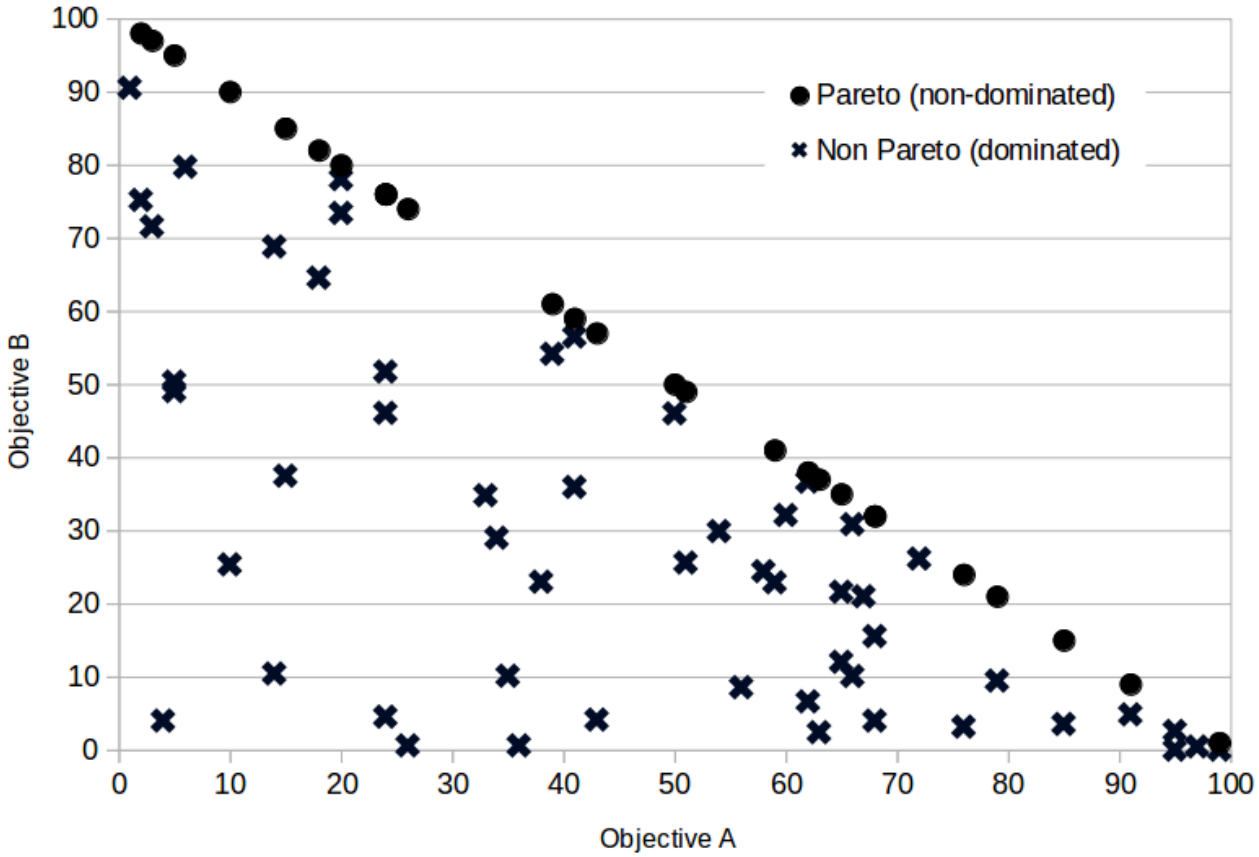


Figure 1: Example of identification of Pareto front (non-dominated) points when comparing two objectives.

The greater the number of objectives on the Pareto Front the lower the chance that a point will be dominated by another. If there is no correlation between objectives and solutions are entirely random then the chance of a single point being dominated by another single point picked at random is $0.5^{n_{obj}}$.

3.2 Algorithm code and data

The algorithm code and data used may be downloaded from.

https://github.com/MichaelAllen1966/1805_thrombectomy_location

(Note: A general version of this algorithm is also available here: https://github.com/MichaelAllen1966/1802_acute_healthcare_location_algorithm)

The algorith was run in two stages:

1) Identify best locations of thrombolysis/thrombectomy centres (this is all that is neeed for the mothership model).

2) For drip'n'ship models, fix the location of thrombectomy centres and identify best locations of thrombolysis-only centres.

Initial algoithm runs focussed on a reduced criteria set of reduce the number of hospitals used, reduce average distance, minimise the the nunebr of admissions to the smallest unit, maximise the size of the smallest unit. This limited approach has two advantages: 1) it finds solutions that are

independnent of specific taregts for distance or admission numbers, and 2) it provdes more praid progress of the algorithm. This progressive extension of objectives is an established general methodology for genetic algorithms [2]. After the initial limited run is complete the numebr of objectives is enlarged (though the algorithm allows for three different target travel times, the next expansion used only one target travel time, with a final run of the algorthm expandinf solutions based on all objectives).

4 Genetic algorithms

Genetic algorithms manage a population of individuals encoded as vectors through a given number of generations. At each generation, 'good' parents are selected from the population according to their fitness (any measure of superiority over other potential parents). Parents are then combined, using a cross-over operator, to create children which are finally mutated. Genetic algorithms differ in the parent selection process, in the cross-over and mutation processes, and in the way the population is archived.

4.1 Representation

Solutions are coded as binary string of genes with either 1 for an open location or 0 for closed. For instance, 001011 would be six genes that represent hospitals 3, 5 and 6 being open and 1, 2 and 4 being closed. In this study, vectors represent the 127 hospitals (SSNAP acute admitting stroke units).

4.2 Selection

The selection operator chooses a part of the population to become parents. The better individuals in terms of objective values are more likely to become parents. The selection probability can be proportionate to fitness by roulette-wheel sampling[3] or stochastic universal sampling[4]. The sigma scaling method normalises the fitness by its variance in the population, so that the individuals with the highest fitness always have a higher probability than others to produce children. However, these approaches focus on exploitation of existing population rather than exploration of the decision space and they can lead to premature convergence.

Other selection methods rely on ranking rather than fitness value. With ranking selection, individuals are ranked according to their fitness and their probability to become parents is function of their rank[5]. Similarly, the tournament selection creates random pairs of individuals and keeps the one with the highest fitness value with a given probability[3]. Such methods allow the algorithm to keep some individuals with low fitness values (with the advantage of keeping a broader gene pool).

Finally, the Boltzmann selection[6] controls the selection rate via a temperature. At the beginning, all individuals have a similar probability to be selected. As the temperature decreases, the selection focusses on high-fitness individuals.

4.3 Cross-over

The cross-over is the process which exchanges genes from parents to create new children. The simplest option is the single-point cross-over which selects one locus and exchanges the blocks of parents before and after that locus. For instance, a crossover at point five would perform the following:

Parent A: 1 1 1 1 1 1

Parent B: 0 0 0 0 0 0

Child A: = 1 1 1 1 0 0

Child B: 0 0 0 0 1 1

The choice of the single-point location can be made by a uniform distribution. In the case of binary vectors, the single-point cross-over is less likely to exchange the endpoints of vectors [2]. To reduce this effect, the cross-over can rely on two or more exchange points.

4.4 Mutation

Mutation changes the gene value of each locus, with a very small probability for each individual each generation. According to [7], the mutation process avoids the loss of diversity in the population.

4.5 Archive

Genetic algorithms also vary by the way solutions are archived and if the population size is variable. The simple option is to keep only children. However, it assumes that children are better than parents which are lost. Several methods build an archive which is union of parents and children. If the population size is variable, an option is to keep the Pareto Front of this archive. However, the size of this Pareto Front can increase dramatically, in particular with many objective functions. Then, individuals from the archive are ranked, based on their Pareto dominance and another metric. NSGA-II [8] and SPEA2 [9] both rank individuals by combining dominance and spread metric in order to maximise population diversity.

4.6 The NSGA-II method

In NSGA-II [8] the archive and the new population are merged and all individuals are ranked according to a two-step mechanism. In the first step, the merged population is split into layers of non-dominated fronts, the first layer being the Pareto Front (the second layer being the next Pareto Front after removal of the first layer). In the second step, the spread of the population is measured by the crowding distance which gives the distance from an individual to its nearest neighbour. To keep the size of the population constant, a given number of individuals is selected from the merged population, preferably from the upper layers and with the largest crowding distance.

NSGA-II has the advantage to keep not only optimal solutions but also near-optimal solutions in lower layers. However, to do so, the population must be large enough. The second advantage is to provide a diverse population in terms of score values, thanks to the crowding distance ranking.

The NSGA-II was chosen for this study after a pilot comparison with SPEA2[9], MOEAD[10], and HypE[11] which showed that NSGA-II provided similar objective performances with a more diverse population.

4.7 Convergence indicator

Population diversity can be monitored using average Hamming distance. The Hamming distance between any two solutions is the proportion of genes that are different. Average Hamming distance

is the mean Hamming distances for all pairwise comparisons in the population (after first Pareto Front selection).

4.8 Description of our genetic algorithm

The code contains a bespoke implementation of a genetic algorithm, based on NSGA-II[8]. Our method evolves solutions based on multiple objectives, but without any weighting of objectives. In each generation, the Pareto Front of non-dominated solutions is identified. Larger populations may be selected by picking subsequent Pareto Fronts (re-evaluation the Pareto Front after removal of the previous Pareto Front identified). The population size is maintained in the interval $\left[P_{min}; P_{max}\right]$.

The steps of the algorithm are:

1) Identify which combination of objectives to use for selection in algorithm (may be from 2 objectives to all objectives).

2) Set up initial population of solutions (a typical starting population is 10,000 solutions).

i) Randomly choose number of hospitals to open in each solution.

ii) Randomly assign open hospitals.

iii) A library of solutions may be imported instead of, or in addition to, a random population of solutions.

iv) Non-unique solutions are removed.

3) Breed solutions:

i) Choose pairs of solutions at random from the population.

While NSGA-II selects parents with the tournament method based on weighted criteria, our method selects parents randomly to avoid weighting any objective.

ii) Select a single crossover point at random within the solution binary string.

iii) Apply the cross-over operator to produce children.

iv) Randomly mutate children with a probability per element of 0.002.

v) Combine parents and children into a new population.

vi) Remove non-unique solutions and any solutions where all hospitals are closed.

4) Calculate the performance of all solutions against the objectives used for selection.

5) Identify all non-dominated (Pareto Front) solutions

i) If the number of selected solutions is greater than the maximum permitted population size then reduce the number of solutions by either

(1) picking the required number of solutions at random, or

(2) pick two solutions at random and use tournament selection based on crowding distance

ii) If the number of selected solutions is lower than the target population then remove the previously selected non-dominated solutions and repeat the Pareto selection until sufficient solutions have been identified.

6) Repeat steps 3-5 until the maximum number of generations is reached or the algorithm is stopped by another indicator:

i) Stop the algorithm when there is a change of <0.001 in average Hamming distance across 5 generations.

Note: The minimum and maximum number of solutions to pass on to the next generation may be the same number to keep solution size constant. Alternatively, a range of population size may be acceptable (e.g. a minimum number of 1,000 solutions may be chosen, but a maximum number of 5,000 solutions may be permitted. In this case Pareto selection is repeated until at least 1,000 solutions have been selected, but restriction on the number of solutions only occurs if the number of solutions chosen exceeds 5,000).

The time taken to reach convergence depended on the the number of objectives in the Pareto Front. Typical populations sizes and run times were:

- For 3-4 objectives: population sizes of 2,500 to 5,000 were used. Typical run time to convergence on a single core of a 2GHz processor was 48hrs.
- For 8-12 objectives: population sizes of 5,00 to 10,000 were used. Typical run time to convergence on a single core of a 2GHz processor was 4-7 days.

Note: algorithms may be speeded up by restricting solutions to a smaller range of acceptable number of hospitals (strict filters may be introduced into the algorithm to remove unacceptable solutions before identifying the Pareto Front).

5 References

1 Zhou A, Qu B-Y, Li H, *et al.* Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm Evol Comput* 2011;**1**:32–49. doi:10.1016/j.swevo.2011.03.001

2 Mitchell M. An introduction to genetic algorithms. *Comput Math with Appl* 1996;**32**:133. doi:10.1016/S0898-1221(96)90227-8

3 Goldberg DE. *Genetic Algorithms in Search, Optimization, and Machine Learning*. 1989. doi:10.1007/s10589-009-9261-6

4 Baker JE. Reducing bias and inefficiency in the selection algorithm. *Proc Second Int Conf Genet Algorithms Genet algorithms their Appl* 1987;:14–21.

5 Baker JE. Adaptive selection methods for genetic algorithms. *Proc an Int Conf Genet Algorithms their Appl* 1985;:101–11.

6 Maza MDA, Tidor B. An analysis of selection procedures with particular attention paid to proportional and Boltzmann selectiono Title. *Proc 5th Int Conf Genet Algorithms* 1993;:124–31.

7 Holland JH. *Adaptation in Natural and Artificial Systems: An introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. 1975. doi:10.1137/1018105

8 Deb K, Pratap A, Agarwal S, *et al.* A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 2002;**6**:182–97. doi:10.1109/4235.996017

9 Zitzler E, Laumanns M, Thiele L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. *Evol Methods Des Optim Control with Appl to Ind Probl* 2001;:95–100. doi:10.1.1.28.7571

10 Zhang Q, Li H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans Evol Comput* 2007;**11**:712–31. doi:10.1109/TEVC.2007.892759

11 Bader J, Zitzler E. HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization. *Evol Comput* 2011;**19**:45–76. doi:10.1162/EVCO_a_00009

Appendix 2: Example scenarios

Four scenarios were mapped and analysed in more detail as examples. These are not intended to describe definitive solutions, but were selected as examples of good solutions:

I. Mothership model with 24 comprehensive stroke centres (CSC) providing intravenous thrombolysis (IVT) and endovascular thrombectomy (ET), where all patients attend their closest CSC. Locations based on the existing 24 neuroscience centres.

II. Mothership model with 30 CSCs providing IVT and ET, where patients attend closest centre. Locations based on using all 18 neuroscience centres outside of London, 4 London neuroscience centres, and 8 further centres that maximise the proportion of patients within 45 minutes of a CSC while keeping all admissions to hospitals above 1,500.

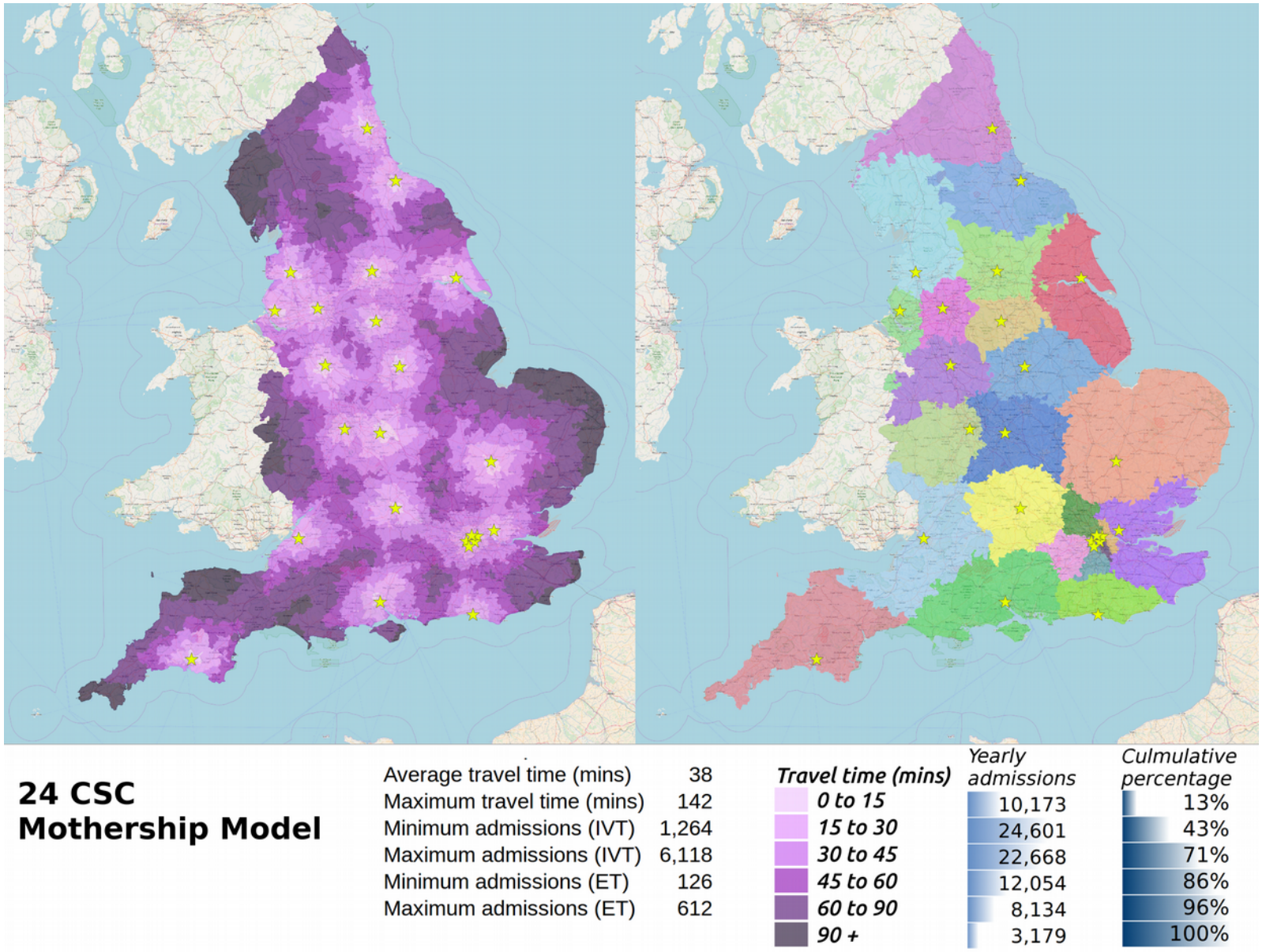
III. Drip n ship model with 30 CSCs as in scenario II. The location of 50 additional Hyper Acute Stroke Units (HASU) were selected in order to maximise the proportion of patients within 30 minutes of a HASU while keeping all the first admissions to hospitals in the range 600 to 2,00 per year. Patients first attend their closest unit regardless of type with onwards travel to CSC if patient requires endovascualr thrombectomy (ET). Show a map for IVT and for ET

IV. Drip n ship model with 30 CSCs and 50 HASUs as in scenario III, but with a 15 minute allowable delay to IVT to first attend a CSC in favour of a HASU, with onwards travel to CSC if patient requires endovascualr thrombectomy (ET).

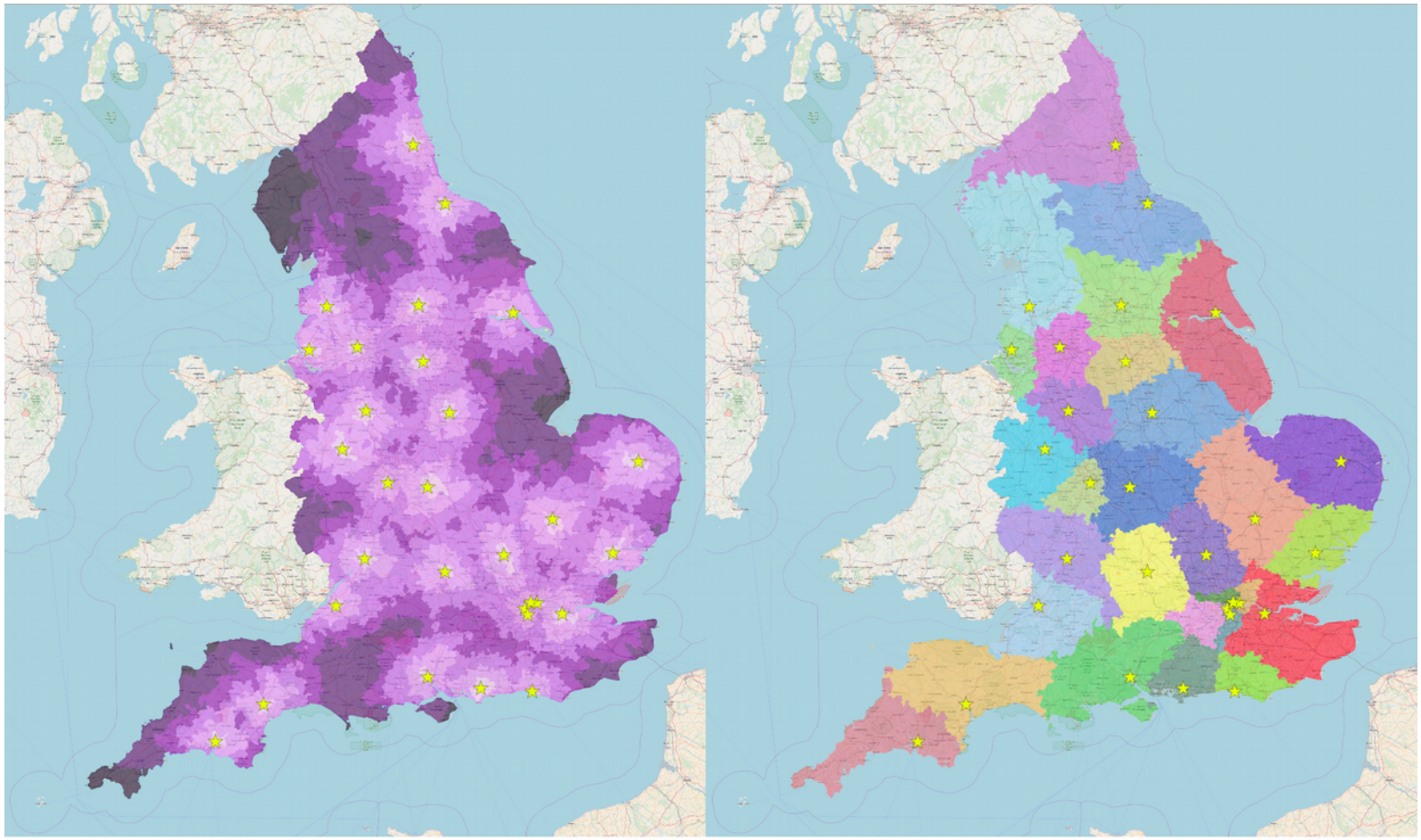
A summary of the performance of the scenarios is shown below:

	Scenario					
	I	II	III		IV	
	24 CSC	30 CSC	30 CSC + 50 HASU Attend closest unit		30 CSC + 50 HASU 15 min allowable delay for CSC	
			IVT	ET	IVT	ET
Number of hospitals	24	30	80	30	80	30
Average travel time (mins)	38	32	22	79	24	54
Maximum travel time (mins)	142	142	99	237	104	237
Minimum admissions (IVT)	1,264	1,532	601	n/a	32	n/a
Maximum admissions (IVT)	6,118	5,723	1,879	n/a	4,320	n/a
Minimum admissions (ET)	126	153	n/a	98	n/a	128
Maximum admissions (ET)	612	572	n/a	690	n/a	603
Patients within 15 mins (%)	13%	14%	34%	13%	28%	14%
Patients within 30 mins (%)	43%	52%	80%	32%	73%	48%
Patients within 45 mins (%)	71%	82%	95%	39%	94%	65%
Patients within 60 mins (%)	86%	93%	98%	40%	98%	69%
Patients within 90 mins (%)	96%	99%	100%	43%	100%	71%
Patients within 120 mins (%)	100%	100%	100%	80%	100%	89%
Patients within 150 mins (%)	100%	100%	100%	94%	100%	96%

Map 1. Scenario 1 - Mothership model with 24 CSC: patient travel times and centre catchment areas for patients attending their closest centre.



Map 2. Scenario 2 - Mothership model with 30 CSC: patient travel times and centre catchment areas for patients attending their closest centre.

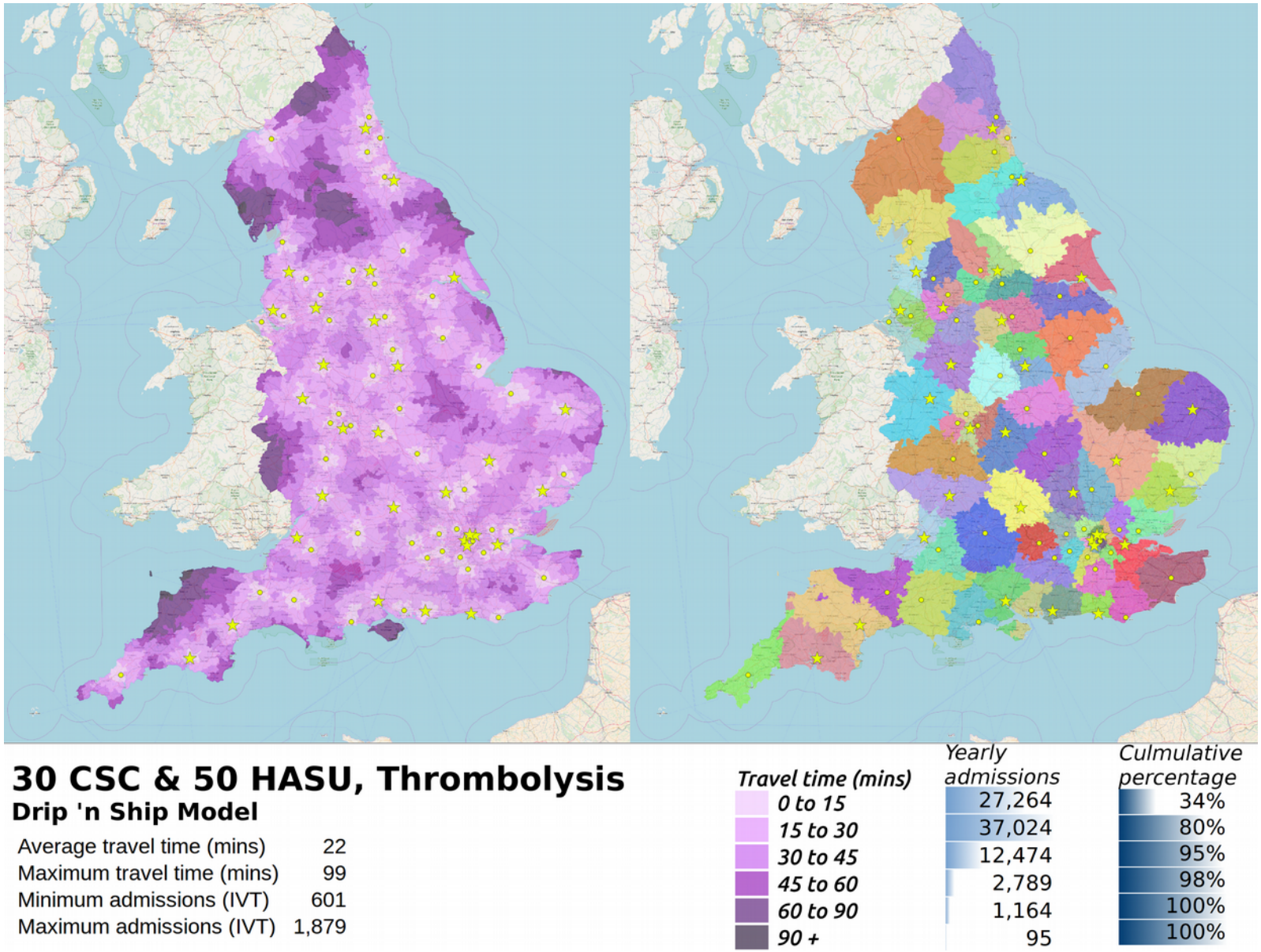


30 CSC Mothership Model

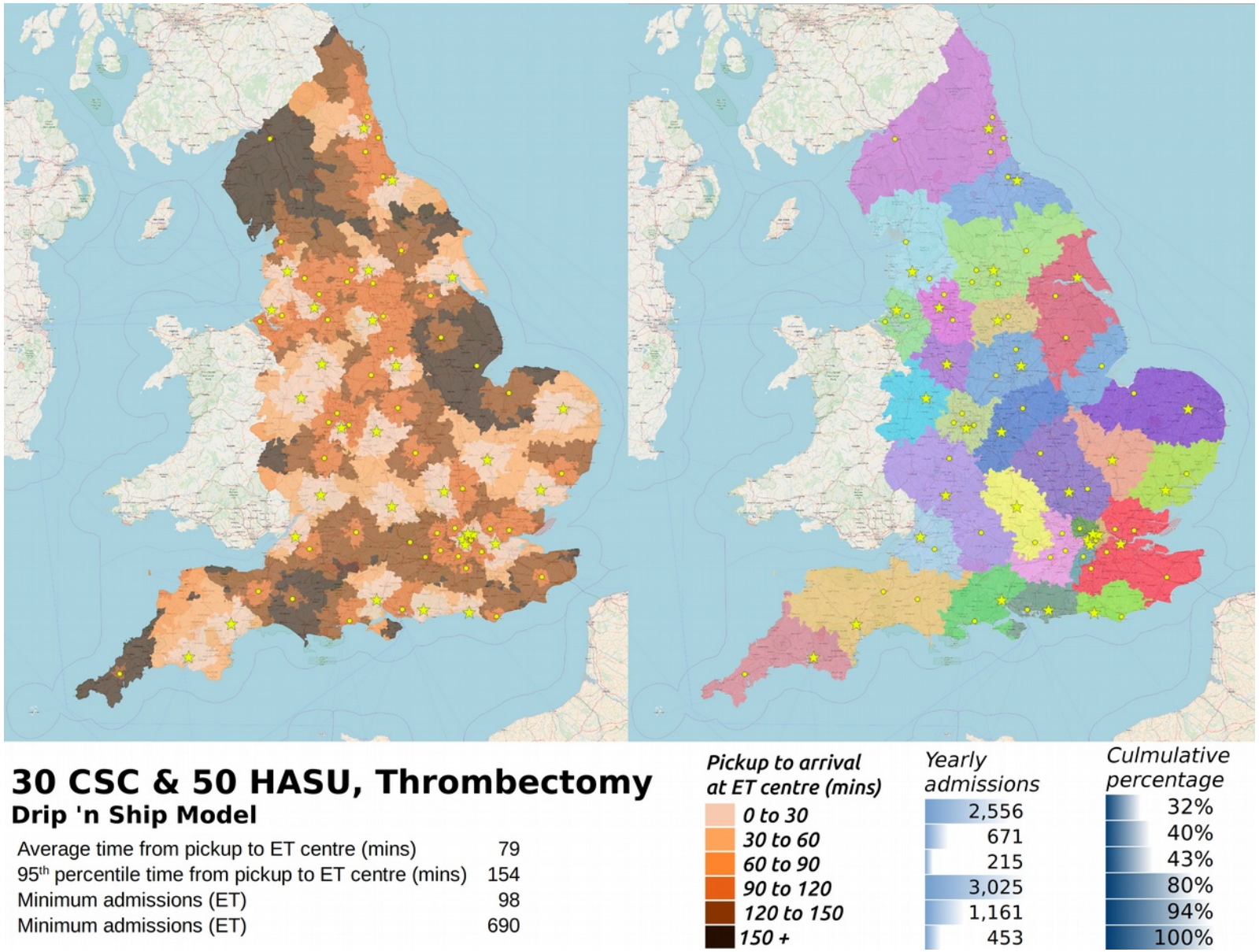
Average travel time (mins)	32
Maximum travel time (mins)	142
Minimum admissions (IVT)	1,532
Maximum admissions (IVT)	5,723
Minimum admissions (ET)	153
Maximum admissions (ET)	572

Travel time (mins)	*Yearly admissions*	*Culmulative percentage*
0 to 15	11,012	14%
15 to 30	31,106	52%
30 to 45	24,431	82%
45 to 60	8,479	93%
60 to 90	4,646	99%
90 +	1,136	100%

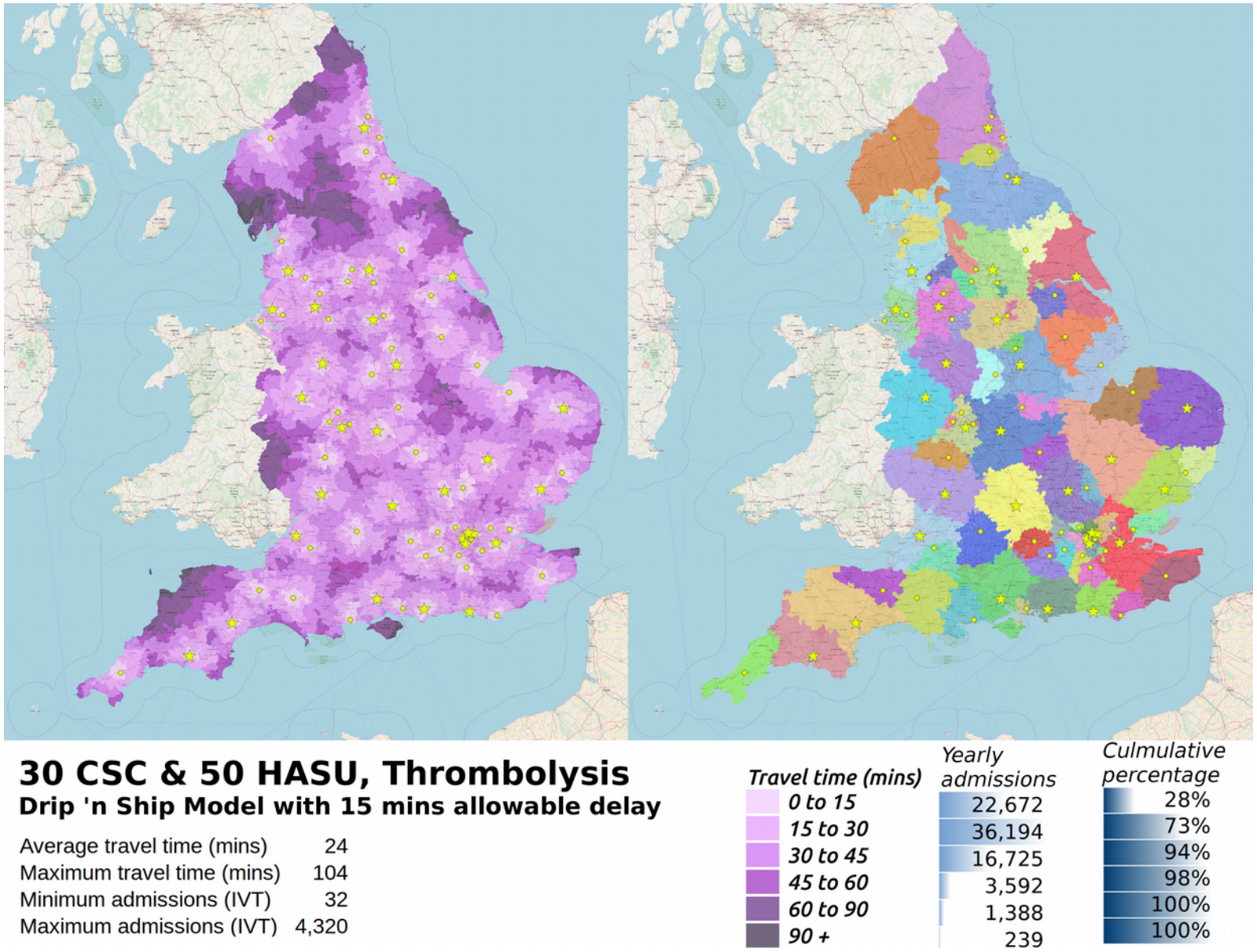
Map 3. Scenario 3 - Drip n ship model with 30 CSC and 50 HASU: patient travel times and centre catchment areas for patients attending first admitted centre (for IVT).



Map 4. Scenario 3 - Drip n ship model with 30 CSC and 50 HASU: patient travel times (including transfer & related delays) and CSC catchment areas for patients attending CSC for ET. Yellow stars show CSC locations, yellow dots show HASU locations.



Map 5. Scenario 4 - Drip n ship model with 30 CSC and 50 HASU with a 15 minute allowable IVT delay: patient travel times and centre catchment areas for patients attending first admitted centre (for IVT).



Map 6. Scenario 4 - Drip n ship model with 30 CSC and 50 HASU with 15 minute allowable IVT delay: patient travel times (including transfer & related delays) and CSC catchment areas for patients attending centre for ET. Yellow stars show CSC locations, yellow dots show HASU locations.

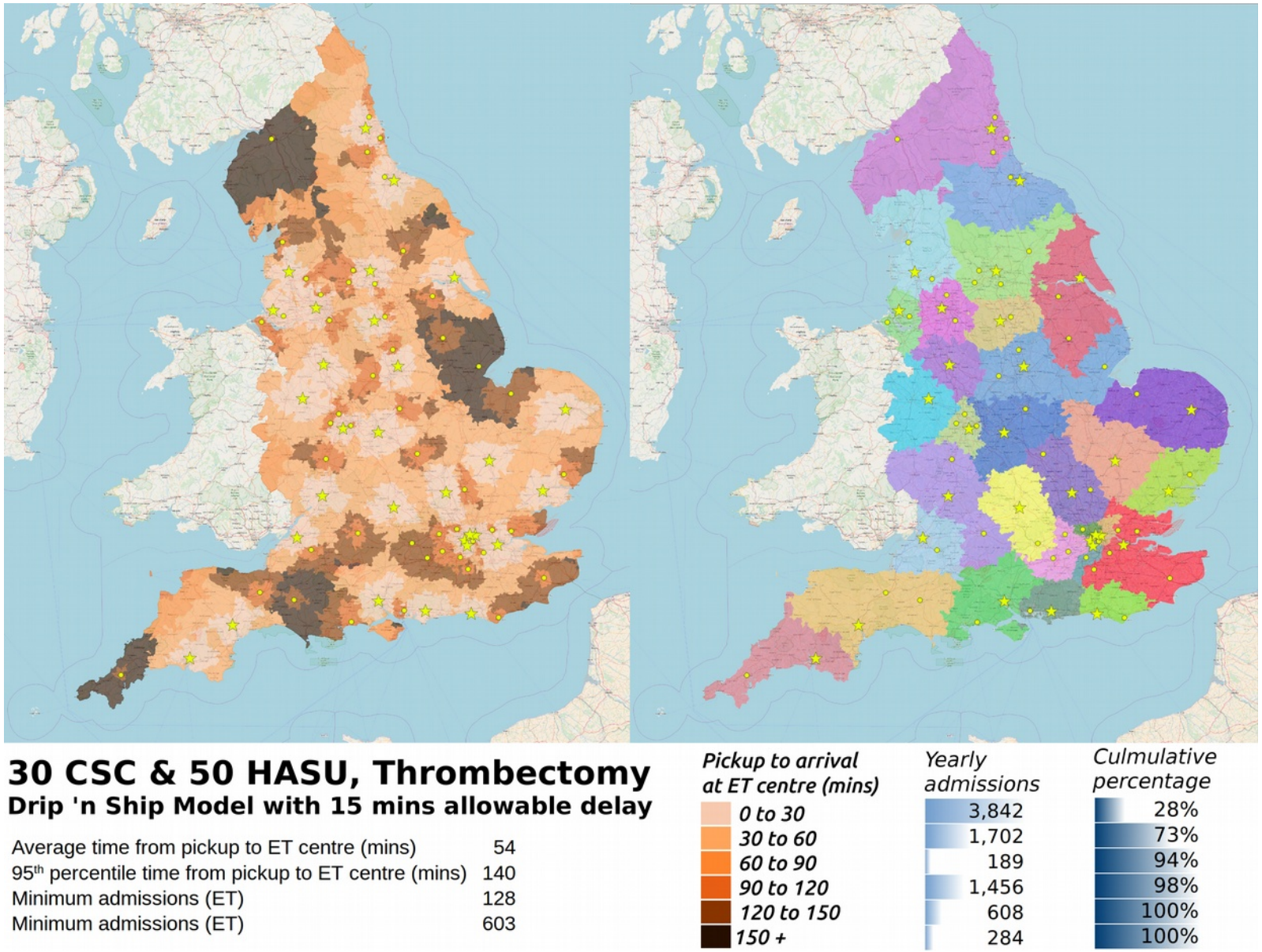


Figure 1. Arrival times to intravenous thrombolysis (IVT) and endovascular thrombectomy (ET) centres. The left panel shows percentile travel times to arrival at first admitting hospital (which is able to provide IVT if required). The right panel shows percentile travel times to an ET-capable centre. Four example scenarios are shown: I) 24 comprehensive stroke centres (CSC) providing IVT and ET where all patients attend closest centre, II) as (I) with 30 CSCs, III) 30 CSC and 50 hyper acute stroke centres (HASU) providing IVT-only, where all patents first attend closest centre with onwards travel to CSC if patient requires endovascular thrombectomy (ET), IV) as (III) but with an allowable IVT delay of 15 minutes.

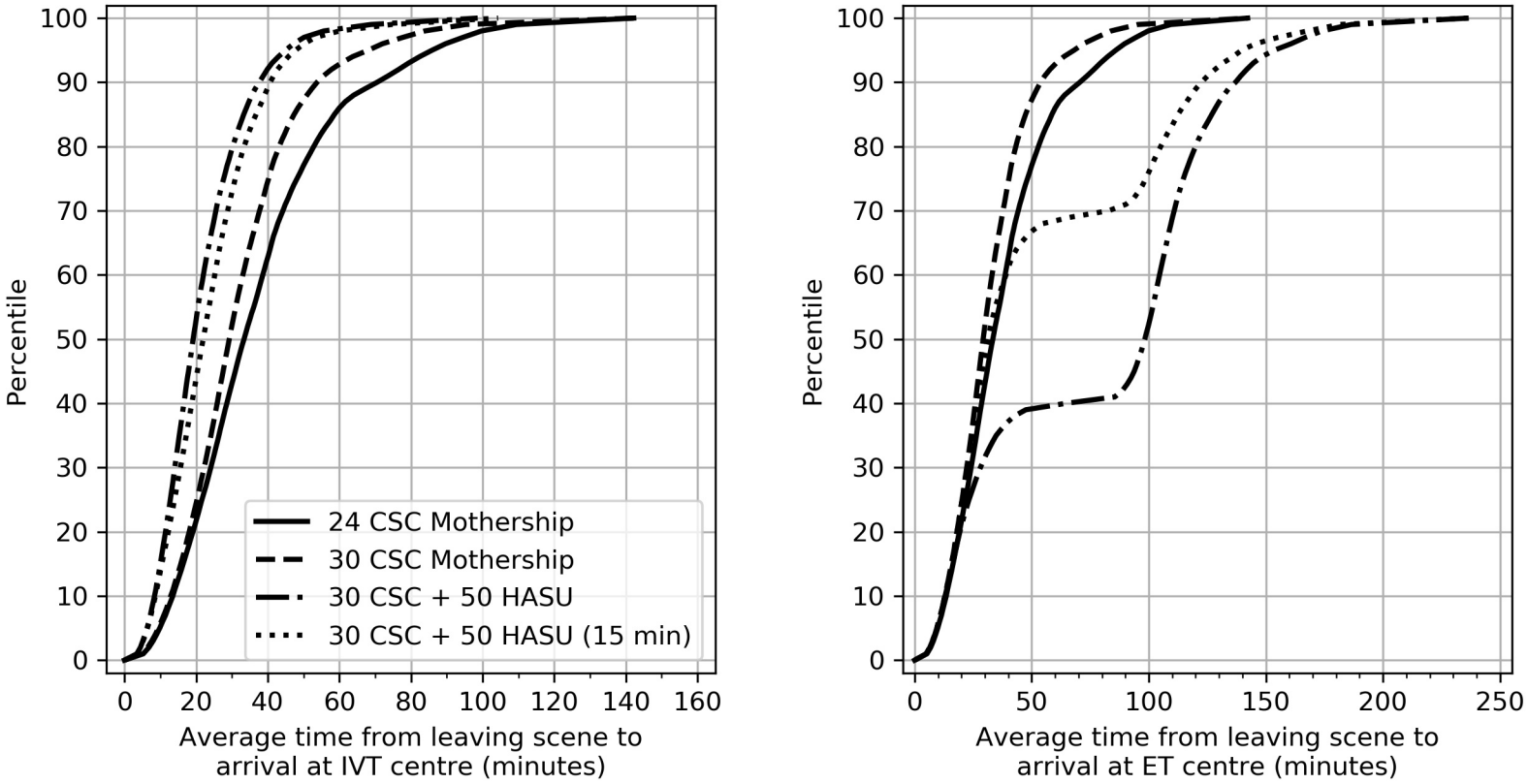


Figure 2. Hospital admissions per year at first admitting hospital (top panel) and ET centre (bottom panel). Four examples scenarios are shown: I) 24 comprehensive stroke centres (CSC) providing IVT and ET where all patients attend closest centre, II) as (I) with 30 CSCs, III) 30 CSC and 50 hyper acute stroke centres (HASU) providing IVT-only, where all patents first attend closest centre with onwards travel to CSC if patient requires endovascular thrombectomy (ET), IV) as (III) but with an allowable IVT delay of 15 minutes.