

1 An Informatics Consult approach for generating clinical evidence for treatment decisions

2

3 Alvina G. Lai^{1,2}, Wai Hoong Chang^{1,2}, Constantinos A. Parisinos¹, Michail Katsoulis¹, Ruth M.
4 Blackburn^{1,2}, Anoop D. Shah^{1,3,12}, Vincent Nguyen¹, Spiros Denaxas^{1,2,3,4}, George Davey Smith^{5,6}, Tom
5 R. Gaunt^{5,6}, Krishnarajah Nirantharakumar^{2,7}, Murray P. Cox⁸, Donall Forde⁹, Folkert
6 Asselbergs^{1,2,3,10,11}, Steve Harris¹², Sylvia Richardson¹³, Reecha Sofat^{1,3}, Richard J.B. Dobson^{1,2,14},
7 Aroon Hingorani^{2,11}, Riyaz Patel¹¹, Jonathan Sterne⁶, Amitava Banerjee^{1,15}, Alastair K. Denniston^{2,16},
8 Simon Ball^{2,16}, Neil J. Sebire^{17,18}, Nigam H. Shah¹⁹, Graham R. Foster²⁰, Bryan Williams^{3,11,12}, Harry
9 Hemingway^{1,2}

10

11

12 ¹Institute of Health Informatics, University College London, London, UK

13 ²Health Data Research UK, UK

14 ³University College London Hospitals NIHR Biomedical Research Centre, London, UK

15 ⁴The Alan Turing Institute, London, UK

16 ⁵Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK

17 ⁶Population Health Sciences. Bristol Medical School, University of Bristol, Bristol, UK

18 ⁷Institute of Applied Health Research, University of Birmingham, UK

19 ⁸Statistics and Bioinformatics Group, School of Fundamental Sciences, Massey University, New Zealand

20 ⁹Public Health Wales, University Hospital of Wales, Cardiff, UK

21 ¹⁰Department of Cardiology, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

22 ¹¹Institute of Cardiovascular Science, University College London, London, UK

23 ¹²University College London Hospitals NHS Trust, London, UK

24 ¹³Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

25 ¹⁴Department of Biostatistics and Health Informatics, King's College London, London, UK

26 ¹⁵Barts Health NHS Trust, The Royal London Hospital, Whitechapel Rd, London, UK

27 ¹⁶University Hospitals Birmingham NHSFT, Birmingham, UK

28 ¹⁷UCL Great Ormond Street Institute of Child Health, London, UK

29 ¹⁸NIHR Great Ormond Street Hospital Biomedical Research Centre, London, UK

30 ¹⁹Department of Medicine, School of Medicine, Stanford University, Stanford California, USA

31 ²⁰Barts Liver Centre, Blizard Institute, Queen Mary University of London, London, UK

32

33 For correspondence: Alvina G. Lai (alvina.lai@ucl.ac.uk)

34 **Abstract**

35 An Informatics Consult has been proposed in which clinicians request novel evidence from large
36 scale health data resources, tailored to the treatment of a specific patient, with return of results in
37 clinical timescales. However, the availability of such consultations is lacking. We seek to provide an
38 Informatics Consult for a situation where a treatment indication and contraindication coexist in the
39 same patient, i.e., anti-coagulation use for stroke prevention in a patient with both atrial fibrillation
40 (AF) and liver cirrhosis. We examined four sources of evidence for the effect of warfarin on stroke
41 risk (efficacy) or all-cause mortality (safety) from: (i) randomised controlled trials (RCTs), (ii) meta-
42 analysis of prior observational studies, (iii) trial emulation (using population electronic health records
43 (N = 3,854,710) and (iv) genetic evidence (Mendelian randomisation). We developed prototype
44 forms to request an Informatics Consult and return of results in electronic health record systems. We
45 found 0 RCT reports and 0 trials recruiting for patients with AF and cirrhosis. We found broad
46 concordance across the three new sources of evidence we generated. Meta-analysis of prior
47 observational studies showed that warfarin use was associated with lower stroke risk (hazard ratio
48 [HR] = 0.71). In a target trial emulation, warfarin was associated with lower all-cause mortality (HR
49 = 0.61) and ischaemic stroke (HR = 0.27). Mendelian randomisation served as a drug target
50 validation where we found that lower levels of vitamin K1 (warfarin is a vitamin K1 antagonist) are
51 associated with lower stroke risk. A pilot survey with an independent sample of 34 clinicians revealed
52 that 85% of clinicians found information on prognosis useful and that 79% thought that they should
53 have access to the Informatics Consult as a service within their healthcare systems. We identified
54 candidate steps for automation to scale evidence generation and to accelerate the return of results
55 within clinical timescales.

56

57 **Introduction**

58 Evidence informing treatment decisions traditionally takes years to generate and leaves many
59 clinical uncertainties unaddressed[1]. Especially in patients with two or more conditions
60 (multimorbidity), it has been hard to generate evidence tailored to ‘patients like me’ and embed this
61 evidence in clinical decision making using electronic health records. There are thousands of clinical
62 practice recommendations, but only a small proportion (15-20%) of these recommendations are
63 supported by level A (trial) evidence[1–6]. A systematic review of trial registration records found that
64 79% of randomised controlled trials (RCTs) excluded patients with concomitant chronic
65 conditions[6]. The United States Food and Drug Administration, Medicines Healthcare Regulatory
66 Authority and European Medicines Agency are increasingly recognising the role of real-world
67 evidence[7] but guidance thus far has not considered its near real-time generation.

68
69 The Informatics Consult concept has been proposed[8–11] to produce on-demand evidence in which
70 clinicians request novel evidence based on the care of prior patients to inform the treatment of a
71 specific patient, with return of results in decision-relevant clinical timescales. For example, a
72 hepatologist seeing a patient with cirrhosis in the clinic learns that they have developed atrial
73 fibrillation. What evidence is, or could rapidly be, available to inform a decision on anti-coagulation
74 for stroke prevention? This is an example of a treatment indication and a treatment contra-indication
75 coexisting in the same patient. Initial experience of the Informatics Consult from Stanford University,
76 has focussed on questions of prognosis (rather than treatment decisions) and highlighted analytical
77 and scaling challenges in returning results in clinical timescales[11]. However, demonstrators of the
78 Informatics Consult for treatment decisions are lacking.

79
80 Our objective was to demonstrate proof of concept of the Informatics Consult in clinically actionable
81 timescales using the atrial fibrillation-cirrhosis-warfarin example. Specifically, we sought to (i)
82 develop prototype electronic health record forms for requesting an Informatics Consult and return of
83 results, (ii) generate four sources of evidence for an Informatics Consult (evaluate available RCT
84 evidence, meta-analyse prior observational studies, emulate a target trial using electronic health
85 records, and Mendelian randomisation[12,13]), (iii) for each form of evidence, to identify steps
86 necessary for automation to accelerate evidence generation and return of results to clinicians, and
87 scale across multiple exemplars (Table S1) and (iv) explore clinician acceptability of the Informatics
88 Consult.

89

90 **Methods**

91 ***Developing prototype electronic health record (EHR) request form and report form***

92 An Informatics Consult is triggered by a request made by a clinician from the electronic health
93 records (EHR) for multiple novel sources of evidence. We developed prototypes for the request form
94 and the report form in consultation with the Chief Clinical Research Informatics Officer and clinicians
95 (cardiologists, hepatologists and clinical pharmacologists). We illustrate possible forms based on an
96 EHR platform, but similar design principles are relevant in EHRs from other vendors, and in other
97 settings, including primary care.

98

99 ***Retrieving information on currently recruiting trials***

100 We searched for previously reported and currently recruiting trials on anticoagulants in patients with
101 atrial fibrillation, with stroke or mortality in the primary outcome using the ClinicalTrials.gov registry.
102 We then used the 8-digit National Clinical Trial numbers to retrieve detailed information on inclusion
103 and exclusion criteria for each trial to identify if any trials included patients with both atrial fibrillation
104 and cirrhosis.

105

106 ***Meta-analysis of prior observational studies***

107 Study identification: We searched PubMed for peer-reviewed articles using the keywords
108 “antithrombotic”, “anticoagulant”, “warfarin”, “cirrhosis” and “atrial fibrillation”. We considered eligible
109 studies as those reporting the effects of anticoagulation therapy in patients with both liver cirrhosis
110 and atrial fibrillation. We excluded reviews, single case reports, editorials and small case series (<
111 10 cases). Data extraction: We extracted the following variables: author, setting, eligibility criteria,
112 number of patients with atrial fibrillation and cirrhosis, number of patients in treated and untreated
113 groups and summary measures. Analyses were performed following PRISMA guidelines. Outcomes
114 of interest were mortality and ischaemic stroke. Statistical analysis. A meta-analysis of associations
115 was performed by pooling hazard ratios (HRs) or odds ratios (ORs) depending on data availability
116 from observational studies using DerSimonian and Laird random-effects models. We also performed
117 leave-one-out sensitivity analyses.

118

119 ***Target trial emulation***

120 We used population-based EHRs to perform a target trial emulation, which is the application of
121 design principles from RCTs to inform analyses on observational data[14,15]. We obtained
122 informational governance approval from the Medicines Healthcare Regulatory Authority (UK)
123 Independent Scientific Advisory Committee (20_078R) to analyse the Clinical Practice Research
124 Datalink (CPRD) linked to secondary care Hospital Episode Statistics and the Office for National
125 Statistic death registration. The study population was 3,854,710 adults aged ≥ 30 years. Phenotype
126 definitions for atrial fibrillation, cirrhosis and other conditions included as baseline covariates as well
127 as definitions for prescriptions are available at <https://caliberresearch.org/portal> and have previously

128 been validated[16,17]. Phenotypes for primary care records were generated using Read clinical
129 terminology (version 2). Phenotypes for secondary care records were generated using ICD-10 terms.
130

131 We developed a target trial protocol where eligibility criteria, treatment assignment, treatment
132 strategy, follow-up period, causal contrast and statistical analyses were specified. Each component
133 of the trial protocol is matched as closely as possible to the design of a randomised trial with minor
134 modifications to accommodate the use of observational data. We employed the intention-to-treat
135 effect as a causal contrast, which was warfarin initiation versus no initiation at baseline. To emulate
136 a target trial, we ensured that individuals are classified as warfarin initiators versus non-initiators at
137 baseline (i.e., using baseline information to assign baseline treatment status). To perform the
138 intention-to-treat analysis, we assigned individuals to the initiator group if they use warfarin within 3
139 months of the baseline date. The baseline is defined as the latest date by which a patient has both
140 cirrhosis and atrial fibrillation given that all eligibility criteria are met. As we were interested in
141 assessing the effects of warfarin use on stroke, we have also excluded prevalent cases of ischaemic
142 stroke. Individuals were followed until the development of an outcome of interest, which were all-
143 cause mortality and incident ischaemic stroke. Propensity score matching (PSM) analyses were
144 performed to reduce bias by matching the warfarin initiator and non-initiator groups. PSM was
145 performed using the nearest-neighbour matching method (a 1:3 match was performed where
146 possible) with a calliper width of 0.2 of the standard deviation of the logit of the propensity score.
147 The PSM cohort was subjected to analyses of all-cause mortality and incident stroke using the
148 Kaplan-Meier and logrank test method and the Cox proportional hazard regression model. As
149 patients and clinicians would be interested in understanding risk in specific demographic categories,
150 we performed subgroup analyses for all-cause mortality in patients aged ≤ 65 , aged > 65 , men,
151 women and in patients with normal international normalise ratio (INR) measurements.

152
153 ***Genetic evidence: Two-sample Mendelian Randomisation (MR)***

154 As an example of drug target validation in the general population, we performed MR to investigate
155 the causal relationship between warfarin use and stroke risk. Vitamin K1 (phylloquinone) is a central
156 component in the production of blood coagulation factors. Warfarin (a vitamin K antagonist) inhibits
157 the activity of vitamin K epoxide reductase to interfere with the recycling of vitamin K and to reduce
158 blood clotting. We considered four single nucleotide polymorphisms (SNPs) that predict circulating
159 phylloquinone (vitamin K1) selected from a genome-wide meta-analysis study on Europeans[18].
160 Four SNPs, all on separate chromosomes, were selected as they had the strongest association with
161 circulating phylloquinone: rs2108622 (chromosome 19), rs2192574 (chromosome 2), rs4645543
162 (chromosome 8) and rs6862071 (chromosome 5). We retrieved genome wide association study
163 (GWAS) summarised data for stroke outcomes from the MEGASTROKE study[19]. MR was
164 performed using the “MendelianRandomisation” package in R[20]. We explored four methods for
165 MR; inverse-variance weighted (IVW), MR-Egger, simple median and weighted median.

166

167 ***Potential for automating the Informatics Consult***

168 We tasked a review panel with expertise in health informatics, epidemiology, evidence synthesis and
169 computer science (drawn from co-authors) with two questions: 1) What are the most time-consuming
170 tasks for each of the 4 streams of evidence identification and generation? (2) What automation
171 opportunities might be important for acceleration and scaling.

172

173 ***Clinician acceptability of the Informatics Consult***

174 After the initial development of the Consult, to gain insights into the acceptability and feasibility of
175 the Informatics Consult, we conducted a pilot survey with an independent sample of 34 clinicians
176 who had not taken part in the research.

177

178 **Results**

179

180 ***Prototype EHR request form***

181 Clinicians wanted the request form (Figure 1) to include a succinct free text statement of the clinical
182 question, auto propagation of the diagnosis combination from the patient's EHR, and suggested
183 structured treatment, efficacy and safety outcomes, based on inputs from the survey. Some
184 clinicians wanted to further specify eligibility criteria for target trial emulation. Interestingly clinicians
185 wanted not only evidence directly related to treatment effectiveness, but also requested national
186 prevalence estimates, prognosis, and current treatment variation.

187

188 ***Prototype EHR report***

189 We provide an overall summary report based on new evidence on warfarin use (lower all-cause
190 mortality and lower stroke risk) in patients with atrial fibrillation and cirrhosis (Figure 2). Through the
191 Consult report, clinicians will have the opportunity to queue patients for RCTs where relevant. We
192 summarise evidence on prevalence, 1-year background mortality risk, meta-analysis of
193 observational studies, target trial emulation results and genetic evidence. The detailed results
194 included in the report for each form of evidence were provided below.

195

196 ***Reported and currently recruiting randomised trials***

197 We did not identify any previous reported RCTs on anticoagulants relevant to patients with cirrhosis.
198 We identified four currently recruiting anticoagulant trials. All four trials reported exclusion criteria
199 related to liver cirrhosis such as contraindication to anticoagulation, i.e., hepatic impairment and
200 elevated liver function tests (Figure 3).

201

202 ***Meta-analysis of observational studies on warfarin use and ischaemic stroke***

203 We identified 142 articles from PubMed, of which only 4 observational studies remained eligible on
204 full-text review, and were included in the meta-analysis[21–24] (Figure 4A). The pooled hazard ratio
205 (HR) of warfarin use in patients with atrial fibrillation and cirrhosis on stroke was 0.71, 95%
206 confidence interval [CI] (CI = 0.39 – 1.29), with high heterogeneity between studies $I^2 = 73\%$ (Figure
207 4B). We also performed the leave-one-out sensitivity analysis with HRs ranging from 0.54 (CI = 0.30
208 - 1.00) to 0.92 (CI = 0.55 - 1.54). Only 1 study reported warfarin use and all-cause mortality and
209 found a lower mortality risk (HR = 0.65; = 0.55 – 0.76)[21].

210

211 ***Target trial emulation using population-based EHRs***

212 Per the target trial protocol, a cohort encompassing 1,022 individuals fulfilling all eligibility criteria
213 was created (initiators = 443; non-initiators = 579). We performed PSM on 22 baseline covariates
214 and generated a matched cohort involving 235 initiators and 526 non-initiators (Figure 5A), baseline
215 patient characteristics before and after PSM are shown. We estimated an intention-to-treat HR for
216 all-cause mortality of 0.61 (CI = 0.49 – 0.76; $p < 0.0001$) comparing warfarin initiators with non-

217 initiators (Figure 5B). Warfarin used was associated with lower risk of ischaemic stroke: HR = 0.27
218 (CI = 0.08 - 0.91, p = 0.034) (Figure 5B). The 1,022 eligible participants for the target trial were
219 categorised into the five subgroups (aged \leq 65, aged $>$ 65, men, women and patients with normal
220 INR), followed by PSM. Intention-to-treat HRs for all-cause mortality comparing warfarin initiators
221 versus non-initiators were as follow: aged \leq 65: HR = 0.62 (0.45 - 0.86, p = 0.0041); aged $>$ 65: HR
222 = 0.61 (0.46 - 0.83, p = 0.0015); men: HR = 0.64 (0.49 - 0.84, p = 0.0014); women: HR = 0.53 (0.37
223 - 0.77, p = 0.00087) and normal INR of $<$ 1.7: HR = 0.62 (0.50 - 0.78, p $<$ 0.0001), where warfarin
224 therapy is associated with lower mortality risk.

225

226 **Genetic evidence**

227 We found no GWAS summary data for vitamin K1 in patients with atrial fibrillation and cirrhosis, but
228 we did find genetic evidence to indicate that warfarin use is associated with reduced stroke risk. For
229 two-sample MR, we used four methods (inverse-variance weighted (IVW), MR-Egger, simple median
230 and weighted median). The IVW analyses, which assumes no pleiotropy, revealed that higher
231 genetically predicted levels of vitamin K1 were associated with a higher risk of any stroke with an
232 odds ratio (OR) of 1.06 (95% CI: 1.00 - 1.11) per Ln-nmol/L increase in vitamin K1 (Figure 6).
233 However, these results were not replicated using methods (simple median, weighted median and
234 MR-Egger) which allow for genetic pleiotropy. When considering stroke subtypes, we observed that
235 higher genetically predicted levels of vitamin K1 were associated with a higher risk of large artery
236 atherosclerotic stroke for 3 out of 4 MR methods: simple median (OR = 1.25 [1.03 - 1.51]); weighted
237 median (OR = 1.25 [1.03 - 1.52]) and IVW (OR = 1.29 [1.11 - 1.50]) (Figure 6).

238

239 **Timely report generation**

240 Generating the evidence and report took one month (two analysts from scratch, without having
241 previous protocols to follow). EHR phenotypes were, however, already created, validated and
242 implemented which sped up the process significantly (<https://portal.caliberresearch.org/>)[17]. Table
243 1 shows additional opportunities for pipelining each of the four streams of evidence identification and
244 generation to return the Consult report. Computable EHR phenotypes and computable clinical trial
245 protocols can be used to automate the process of trial identification and trial recruitment[25,26]. For
246 meta-analyses of observational studies, approaches for semi-automated systematic reviews[27,28],
247 batch extraction of data from articles[29,30] and mapping of SNOMED-CT terms to MeSH
248 descriptors in PubMed[31] can be used in the pipelining process. Automating these steps and
249 performing data extractions from articles in batch mode will enable the return of results within 24
250 hours. The DExtER tool can be used for automatic extraction of EHR databases and automated
251 cohort creation for the target trial emulation process[32]. Recent addition of an analytical module to
252 DExtER enables cohort creation, statistical analyses and results visualisation within short time
253 scales (1-4 hours) for matched cohort studies. Ongoing work will make DExtER return results real-
254 time. Two-sample MR to generate genetic evidence can be automated using the MR-Base

255 platform[33], which returns results within minutes. By using such approaches to automation, we
256 estimate all four streams of new evidence could be generated in 48 hours by one health informatician
257 from the point at which the clinician makes the Consult request.

258

259 ***Clinician survey on the acceptability and feasibility of the Informatics Consult***

260 We surveyed an independent sample of 34 clinicians from eight specialties with results shown in
261 (Figure S2). Results indicated that 79% of clinicians thought that they should have access to the
262 Informatics Consult as a service within their healthcare systems (21% responded 'maybe'). Clinicians
263 found each section of the report useful (or 'maybe useful') as follows: prognosis 85% (12%), a
264 summary of evidence of efficacy and safety 79% (21%) and disease prevalence 68% (18%). Only
265 18% of clinicians thought that they needed to have access to the details of the evidence in the clinic
266 with one clinician stating: "*Multi-disciplinary team meetings might benefit from Informatics Consults,*
267 *this is where difficult cases are discussed and there is time to review newly generated evidence*".
268 When asked whether clinicians found detailed reports on the four sources of evidence useful (or
269 'maybe useful'), responses were as follow: randomised trials: 82% (15%); meta-analysis of prior
270 observational evidence: 76% (18%); target trial emulation: 62% (18%) and genetic evidence: 26%
271 (24%). Clinicians would discuss the Informatics Consult report with their patients 74% (26%
272 answered maybe). Clinicians offered 15 further clinical questions where the Informatics Consult
273 might be of value and made additional comments including: "*Ultimately, we practise defensive*
274 *medicine - would my decision stand up in court based on available data – the Informatics Consult*
275 *should help with that.*" and "*Regulators and guideline developers will require replication and quality*
276 *assurance of evidence generated in clinical timescales*".

277

278 **Discussion**

279 We demonstrate that the Informatics Consult offers a novel paradigm to generate new clinical
280 evidence and return results within clinical timescales. We found that in patients with atrial fibrillation
281 and cirrhosis, initiation of warfarin was common, may be associated with lower all-cause mortality
282 and may be effective in lowering stroke risk. Given the ubiquity of clinical uncertainty where there is
283 little or no evidence, and that current modes of generating new evidence may never be initiated or,
284 if initiated, take years to report, it is imperative to accelerate learning from extant data.

285

286 ***Informatics Consult versus traditional approaches for evidence generation and delivery***

287 The Informatics Consult puts the treating clinician and the patient at the centre of evidence
288 generation. Indeed, in seeking to address a range of questions from the clinician and patient, the
289 Informatics Consult enables simultaneous delivery of evidence from different sources, rather than
290 employing a one-study-one-design-at-a-time approach. The Consult is embedded within EHR
291 system and evidence is generated within clinical timescales—making it a form of an electronic
292 consult, which is increasingly being adopted to seek specialist input[34,35]. Additional information
293 on how the Informatics Consult differs from traditional approaches to evidence generation and use
294 is summarised in Table S1.

295

296 ***Concordance across sources of evidence identified from the Consult***

297 A primary motivation for requesting an Informatics Consult is to understand how a particular
298 treatment influences an outcome to help guide decision making. We show a degree of concordance
299 across four sources of evidence. RCTs demonstrate the effectiveness of oral anticoagulation in
300 stroke prevention in patients with atrial fibrillation: but as we demonstrate reported and currently
301 recruiting RCTs exclude patients with cirrhosis (i.e., patients having ‘contraindications’ for
302 anticoagulants). This suggests that the prospect of ever mounting, or successfully recruiting to, an
303 RCT in patients with AF and cirrhosis is low. Meta-analysis of observational studies and target trial
304 emulation suggest evidence on the potential benefits of warfarin for stroke reduction, suggesting the
305 significant impact on strokes and deaths averted if these patients are treated with anticoagulant
306 therapy. Although we did not find any relevant GWAS summary data, evidence suggests that lower
307 levels of vitamin K1 (target of warfarin) are associated with lower stroke risk which is corroborated
308 by another study[36].

309

310 ***Returning results on prevalence, prognosis and treatment variation***

311 Interestingly, the clinicians involved in this study recommended that information beyond efficacy and
312 safety should be included as options for the clinician to request in the Informatics Consult. By
313 providing information on the prevalence of cirrhosis and atrial fibrillation, we demonstrate that this
314 pair of conditions is not highly prevalent (although not considered rare). Knowing that the health
315 system has data on diagnosis, treatment and outcomes in an estimated 35,000 individuals (scaled

316 up to the population of England) with both conditions, highlights the importance of several areas.
317 First, the importance of accessing and learning from nationwide data at scale; any one clinician may
318 have clinical experience of only a handful of such cases, and it is not feasible or scalable to create
319 registries. Second, the clinician likely has never had access before to population-based,
320 contemporary prognostic information on such patients (1-year mortality: 15%), nor the knowledge
321 that in practice 43% of patients were started on warfarin. We anticipate that providing clinicians and
322 patients with this information may stimulate further questions on generating new Consults on a wider
323 range of prognostic outcomes, and predictors of prognosis.

324

325 ***Need and demand for Informatics Consult***

326 Nearly all clinicians are faced with treatment decisions where limited evidence exists where insights
327 might be gained from analysis of large-scale data on 'patients like me'. But currently, few if any
328 clinicians can request such insights. A clinician from the survey mentioned that evidence from the
329 Informatics Consult is important to help their decision in practicing defensive medicine given that the
330 majority of clinical practice recommendations in professional society guidelines are not supported by
331 RCT evidence. There is limited information on the system-wide frequency of treatment 'clashes'
332 where indication and contraindications coexist in the same patient. This is especially relevant for
333 patients with multimorbidity where evidence from RCTs are limited[6], which may result in individuals
334 being subjected to low-quality recommendations. For example, certain targeted cancer therapy such
335 as angiogenesis inhibitors may cause an increase the prevalence of hypertension during treatment,
336 which highlights the importance of considering the impact of cancer therapy on adverse side effects
337 and cardiotoxicity[37]. Our preliminary analyses demonstrate that 26 in 10,000 women have breast
338 cancer and hypertension; these individuals might benefit from the Informatics Consult.

339

340 ***Developing the Informatics Consult as a service***

341 The majority of clinicians in our survey thought that they should have access to the Informatics
342 Consult as a service as their healthcare systems seek to learn from existing data[38]. We have
343 demonstrated the feasibility with this one example. Scaling to other clinical examples, and a service,
344 requires five inter-related challenges to be addressed in clinical standards setting, implementation
345 and evaluation, access to data, informatics, and knowledge management. First, bodies that define
346 standards of care – including clinical practice guideline developers and regulatory and technology
347 assessment bodies – are already considering real-world evidence generated in conventional
348 timescales[7,39]. In the context of a more rapid generation of evidence, the Informatics Consult
349 raises new questions about replication, open peer review and quality assurance pipelines. Second,
350 development, implementation and evaluation of the impact of the Informatics Consult on clinical
351 decision making in practice are required. Feedback from our survey suggests that Multidisciplinary
352 Team (MDT) meetings, which concentrate on 'difficult' clinical cases may present an opportunity[40].
353 Third, an Informatics Consult service requires approved, immediate access to large scale, clinically

354 detailed updated data, most likely in a trusted research environment. The coronavirus pandemic
355 offers a new precedent for such information governance approvals, which can not foresee the nature
356 of the next clinician's question. Fourth, biomedical knowledge (in EHRs, trial protocols, clinical
357 guidelines) needs to become more computable and interoperable. Design and analysis methods
358 need to be improved, standardized and widely distributed. For example, the Observational Health
359 Data Science and Informatics community offers open-source software implementing many routine
360 analyses methods[41]. It took one month to generate simultaneously the findings on four streams of
361 evidence. Without the Informatics Consult, this evidence may never be generated at all; or may be
362 published at different times over a period of years. A feasible goal, using automation, would be to
363 generate the report in 48 hours. Fifth, there needs to be an open process of making the knowledge
364 available in an Informatics Consult library. As the library builds, Clinicians might make a request for
365 which a previous consult already provides an answer. The Library might also serve as a platform for
366 connecting 'patients like me', registering the frequency of therapeutic dilemmas and potential
367 treatment uncertainties, identifying the need for new RCTs, and informing their design and facilitate
368 targeted recruitment into trials.

369

370 ***Strengths of the study***

371 To the best of our knowledge, this is the first demonstration of Informatics Consult for a treatment
372 decision, triangulation of evidence from meta-analyses of observational studies, target trial emulation
373 using EHR data and MR. The breadth, depth and longevity (long follow-up period) of the EHR data,
374 which links primary care, secondary care and the death registry is an advantage. Another significant
375 strength is engagement with an independent sample of clinicians from multiple specialties to gauge
376 the feasibility and acceptability of the Informatics Consult.

377

378 ***Limitations of the study***

379 Our study has important limitations. First, although we have prototyped EHR request and report
380 forms based on feedback from clinicians, these have not yet been implemented in live clinical
381 systems. Second, as initial proof of concept, we have not assessed bleeding outcomes and newer
382 non-vitamin K oral anticoagulants due to current data access limitations. Third, it remains difficult to
383 assess whether there unmeasured confounding despite employing causal inference methods. It is
384 also possible that some prescription data is missing, which means that patients are incorrectly
385 assigned to the control arms. Future analyses on a more recent and larger dataset involving both
386 CPRD Gold and Aurum[42] would be beneficial. Fourth, a limitation of the MR analyses is that access
387 to summarised GWAS data for stroke outcomes in patients with cirrhosis is limited. The advantages
388 of using publicly available GWAS summarised data are speed and transparency, both of which are
389 essential to the Informatics Consult. Although individual-level data would allow more flexibility to
390 conduct analyses in specific patient subgroups and to select which variables to generate the
391 summarised data for, such analyses could not be returned within a clinical timescale that is not

392 scalable and cannot be fully automated. Fifth, although we have identified potential ways to automate
393 the analytic process, these have not been implemented here.

394

395 **Conclusion**

396 We proposed an Informatics Consult framework to summarize evidence from four sources and have
397 developed a report prototype for answering a treatment question to enable new ways of data-
398 informed decision making in clinical timescales. The Informatics Consult may stimulate a
399 conversation among public, professionals and policymakers about more rapidly realising the benefits
400 of health system learning from 'patients like me'.

401

402 **Acknowledgements**

403 We thank clinical colleagues who have contributed to the pilot survey and who have consented for
404 their names to be acknowledged: Christopher Tomlinson, Upkar Gill, Paul Harrow, Harry Martin,
405 Omer Ahmad, Louise China, Vanessa Taylor, Sai Ambati , Amy Prideaux, Alex May, Ruth Gilbert,
406 Hamish Miller, Bilal Mateen, Rhys Davies , Bu'Hussain Hayee, David Cronin, Katharine Pollock and
407 Philip Oppong.

408

409 **Authors' contributions**

410 Research question: AGL, GRF, HH

411 Funding: AGL, HH

412 Study design and analysis plan: AGL, WHC, CP, HH

413 Preparation of data, including electronic health record phenotyping in the CALIBER portal: AGL,
414 WHC, CP, SD

415 Statistical analysis: AGL, WHC, MK

416 Drafting initial versions of the manuscript: AGL, HH

417 Critical review of early and final versions of the manuscript: All authors

418

419 **Conflicts of interest:**

420 AB has received research funding from AstraZeneca for work unrelated to this research. GRF
421 receives funding from companies that manufacture drugs for hepatitis C virus (AbbVie, Gilead, MSD)
422 and consults for GSK, Arbutus and Shionogi in areas unrelated to this research. TRG and GDS have
423 received research funding from GlaxoSmithKline and Biogen for work unrelated to this research.

424

425 **Funding statement:**

426 AGL is supported by funding from the Wellcome Trust (204841/Z/16/Z), National Institute for Health
427 Research (NIHR) University College London Hospitals Biomedical Research Centre
428 (BRC714/HI/RW/101440), NIHR Great Ormond Street Hospital Biomedical Research Centre
429 (19RX02) and the Health Data Research UK Better Care Catalyst Award (CFC0125). MK is funded
430 by the British Heart Foundation (FS/18/5/33319). RMB is supported by a UKRI Innovation Fellowship
431 funded by the Medical Research Council (Grant No: MR/S003797/1). ADS is supported by a
432 postdoctoral fellowship from THIS Institute. RJBD is supported by the NIHR Biomedical Research
433 Centres at South London and Maudsley NHS Foundation Trust (SLAM; IS-BRC1215-20018); Health
434 Data Research UK; UK Research and Innovation (UKRI) London Medical Imaging & Artificial
435 Intelligence Centre for Value Based Healthcare; the BigData@Heart Consortium (Grant No. 116074
436 of the European Union Horizon 2020 programme); the NIHR BRC and Research Informatics Unit at
437 University College London Hospitals; and the NIHR Applied Research Collaboration South London
438 at KCHFT. GDS and TG are funded by the Medical Research Council Integrative Epidemiology Unit
439 at the University of Bristol MC_UU_00011/1&4. HH is an NIHR Senior Investigator and is funded by

440 the NIHR University College London Hospitals Biomedical Research Centre, supported by Health
441 Data Research UK (LOND1).

442 **References**

- 443 1. van Dijk WB, Grobbee DE, de Vries MC, Groenwold RHH, van der Graaf R, Schuit E. A
444 systematic breakdown of the levels of evidence supporting the European Society of Cardiology
445 guidelines. *Eur J Prev Cardiol*. SAGE Publications Sage UK: London, England; 2019;26:1944–52.
- 446 2. Meyer C, Bowers A, Wayant C, Checketts J, Scott J, Musuvathy S, et al. Scientific evidence
447 underlying the American College of Gastroenterology’s clinical practice guidelines. *PLoS One*.
448 Public Library of Science San Francisco, CA USA; 2018;13:e0204720.
- 449 3. Koh C, Zhao X, Samala N, Sakiani S, Liang TJ, Talwalkar JA. AASLD clinical practice
450 guidelines: a critical review of scientific evidence and evolving recommendations. *Hepatology*.
451 Wiley Online Library; 2013;58:2142–52.
- 452 4. Fanaroff AC, Califf RM, Windecker S, Smith SC, Lopes RD. Levels of Evidence Supporting
453 American College of Cardiology/American Heart Association and European Society of Cardiology
454 Guidelines, 2008-2018. *Jama*. American Medical Association; 2019;321:1069–80.
- 455 5. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC. Scientific evidence underlying the
456 ACC/AHA clinical practice guidelines. *Jama*. American Medical Association; 2009;301:831–41.
- 457 6. Du Vaure CB, Dechartres A, Battin C, Ravaud P, Boutron I. Exclusion of patients with
458 concomitant chronic conditions in ongoing randomised controlled trials targeting 10 common
459 chronic conditions and registered at ClinicalTrials. gov: a systematic review of registration details.
460 *BMJ Open*. British Medical Journal Publishing Group; 2016;6.
- 461 7. U.S. Food and Drug Administration. Framework for FDA’S real-world evidence program.
462 <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>.
463 Accessed 16 November 2020.
- 464 8. Gombar S, Callahan A, Califf R, Harrington R, Shah NH. It is time to learn from patients like
465 mine. *npj Digit Med* [Internet]. Springer US; 2019;2:2018–20. Available from:
466 <http://dx.doi.org/10.1038/s41746-019-0091-3>
- 467 9. Longhurst CA, Harrington RA, Shah NH. A “green button” for using aggregate patient data at the
468 point of care. *Health Aff*. 2014;33:1229–35.
- 469 10. Callahan A, Gombar S, Jung K, Steinberg E, Harrington R, Shah NH. Delivering on-demand
470 evidence via an informatics consultation service. :3–5.
- 471 11. Schuler A, Callahan A, Jung K, Shah NH. Performing an informatics consult: methods and
472 challenges. *J Am Coll Radiol*. Elsevier; 2018;15:563–8.
- 473 12. Davies NM, Holmes M V., Davey Smith G. Reading Mendelian randomisation studies: A guide,
474 glossary, and checklist for clinicians. *BMJ*. 2018;362.
- 475 13. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for
476 performing Mendelian randomization investigations. *Wellcome Open Res*. 2020;4:186.
- 477 14. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not
478 available. *Am J Epidemiol*. Oxford University Press; 2016;183:758–64.
- 479 15. Dickerman BA, Hernán MA. Avoidable flaws in observational analyses: an application to statins

- 480 and cancer. *Nat Med* [Internet]. Springer US; 2019;25. Available from:
481 <http://dx.doi.org/10.1038/s41591-019-0597-x>
- 482 16. Kuan V, Denaxas S, Gonzalez-Izquierdo A, Direk K, Bhatti O, Husain S, et al. A chronological
483 map of 308 physical and mental health conditions from 4 million individuals in the English National
484 Health Service. *Lancet Digit Heal*. 2019;1:e63–77.
- 485 17. Denaxas S, Gonzalez-Izquierdo A, Direk K, Fitzpatrick NK, Fatemifar G, Banerjee A, et al. UK
486 phenomics platform for developing and validating electronic health record phenotypes: CALIBER. *J*
487 *Am Med Informatics Assoc*. 2019;
- 488 18. Dashti HS, Shea MK, Smith CE, Tanaka T, Hruby A, Richardson K, et al. Meta-analysis of
489 genome-wide association studies for circulating phylloquinone concentrations. *Am J Clin Nutr*.
490 Oxford University Press; 2014;100:1462–9.
- 491 19. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry
492 genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and
493 stroke subtypes. *Nat Genet*. Nature Publishing Group; 2018;50:524–37.
- 494 20. Yavorska OO, Burgess S. MendelianRandomization: An R package for performing Mendelian
495 randomization analyses using summarized data. *Int J Epidemiol*. 2017;46:1734–9.
- 496 21. Serper M, Weinberg EM, Cohen JB, Reese PP, Taddei TH, Kaplan DE. Mortality and hepatic
497 decompensation in patients with cirrhosis and atrial fibrillation treated with anticoagulation.
498 *Hepatology*. 2020;0–3.
- 499 22. Choi J, Kim J, Shim JH, Kim M, Nam GB. Risks Versus Benefits of Anticoagulation for Atrial
500 Fibrillation in Cirrhotic Patients. *J Cardiovasc Pharmacol*. 2017;70:255–62.
- 501 23. Kuo L, Chao TF, Liu CJ, Lin YJ, Chang SL, Lo LW, et al. Liver Cirrhosis in Patients With Atrial
502 Fibrillation: Would Oral Anticoagulation Have a Net Clinical Benefit for Stroke Prevention? *J Am*
503 *Heart Assoc*. 2017;6.
- 504 24. Lee SJ, Uhm JS, Kim JY, Pak HN, Lee MH, Joung B. The safety and efficacy of vitamin K
505 antagonist in patients with atrial fibrillation and liver cirrhosis. *Int J Cardiol*. 2015;180:185–91.
- 506 25. Ahmad FS, Rickett IM, Hammill BG, Eskenazi L, Robertson HR, Curtis LH, et al. Computable
507 Phenotype Implementation for a National, Multicenter Pragmatic Clinical Trial: Lessons Learned
508 From ADAPTABLE. *Circ Cardiovasc Qual Outcomes*. *Am Heart Assoc*; 2020;CIRCOUTCOMES--
509 119.
- 510 26. Willoughby C, Fridsma D, Chatterjee L, Speakman J, Evans J, Kush R. A standard computable
511 clinical trial protocol: the role of the BRIDG model. *Drug Inf J DIJ/Drug Inf Assoc*. Springer;
512 2007;41:383–92.
- 513 27. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using
514 machine learning tools in research synthesis. *Syst Rev*. *BioMed Central*; 2019;8:163.
- 515 28. Kiritchenko S, De Bruijn B, Carini S, Martin J, Sim I. ExaCT: automatic extraction of clinical trial
516 characteristics from journal publications. *BMC Med Inform Decis Mak*. Springer; 2010;10:56.
- 517 29. Constantin A, Pettifer S, Voronkov A. PDFX: fully-automated PDF-to-XML conversion of

- 518 scientific literature. Proc 2013 ACM Symp Doc Eng. 2013. p. 177–80.
- 519 30. Parizi RM, Guo L, Bian Y, Azmoodeh A, Dehghantanha A, Choo K-KR. CyberPDF: smart and
520 secure coordinate-based automated health PDF data batch extraction. 2018 IEEE/ACM Int Conf
521 Connect Heal Appl Syst Eng Technol. 2018. p. 106–11.
- 522 31. Jacobs AK, Quinn TA, Nelson SJ. Mapping SNOMED-CT concepts to MeSH concepts. AMIA
523 Annu Symp Proc. 2006. p. 965.
- 524 32. Gokhale KM, Chandan JS, Toulis K, Gkoutos G, Tino P, Nirantharakumar K. Data extraction
525 for epidemiological research (DExtER): a novel tool for automated clinical epidemiology studies.
526 Eur J Epidemiol. Springer; 2020;1–14.
- 527 33. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform
528 supports systematic causal inference across the human phenome. Elife. eLife Sciences
529 Publications Limited; 2018;7:e34408.
- 530 34. Vimalananda VG, Gupte G, Seraj SM, Orlander J, Berlowitz D, Fincke BG, et al. Electronic
531 consultations (e-consults) to improve access to specialty care: a systematic review and narrative
532 synthesis. J Telemed Telecare. SAGE Publications Sage UK: London, England; 2015;21:323–30.
- 533 35. Informatics Consultation Service at Stanford. <http://greenbutton.stanford.edu>. Accessed 17
534 December 2020.
- 535 36. Larsson SC, Traylor M, Markus HS. Circulating vitamin K1 levels in relation to ischemic stroke
536 and its subtypes: A mendelian randomization study. Nutrients. 2018;10:1–7.
- 537 37. Mouhayar E, Salahudeen A. Hypertension in cancer patients. Texas Hear Inst J. Texas Heart
538 Institute; 2011;38:263.
- 539 38. Budrionis A, Bellika JG. The learning healthcare system: where are we now? A systematic
540 review. J Biomed Inform. Elsevier; 2016;64:87–92.
- 541 39. European Medicines Agency. Guideline on registry-based studies. Accessed 15 December
542 2020. <https://www.ema.europa.eu/en/guideline-registry-based-studies>.
- 543 40. National Guideline Centre (UK). Emergency and acute medical care in over 16s: service
544 delivery and organisation. London: National Institute for Health and Care Excellence (UK); 2018
545 Mar. (NICE Guideline, No. 94.) Chapter 29, Multidisciplinary team meeting.
- 546 41. Observational Health Data Sciences and Informatics Methods Library.
547 <https://www.ohdsi.org/methods-library/>. Accessed 2020-12-07.
- 548 42. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile:
549 Clinical Practice Research Datalink (CPRD) Aurum. Int J Epidemiol. 2019;
550

Figure 1. Electronic health record Informatics Consult request form prototype

Lockhart, Tifa

Female, 46 y.o., 04/05/1974
MRN: 18362946
Arrival date: 06/10/2020

Bed: 5B24

Consultant: Strife, Mandy

Treatment Team: Strife, Mandy; Chen, Liyi

Allergies: No known allergies

Patient Class: Outpatient

Current weight: 74kg

BMI: 28.5

Blood Type: A-positive

Chart Review

Results Review

Medications

Immunisations

Diagnoses

Orders

Results

Informatics Consult Request

Informatics Consult Report

Submit consult

Cancel

Informatics Consult Request Form:

[Informatics Consult service: Who we are and what we do? \(link\)](#)

Clinical question:

Please describe the treatment uncertainty that you would like the Consult to address here. Please include any eligibility criteria.

Which diagnoses to include:

- + Atrial fibrillation
- + Cirrhosis

Please select

What treatment decision?

- + Warfarin

What outcomes?

- + *Safety:* All-cause mortality
- + *Efficacy:* Ischaemic stroke

What evidence do you want?

- + Evidence for treatment effectiveness
- + Other evidence:
 - Prevalence
 - Prognosis
 - Treatment variation

When do you want this evidence?

I would like to see the patient and make a decision within:

- + 48 hours
- + 1 week
- + 1 month



Figure 2. Electronic health record Informatics Consult report prototype.

Lockhart, Tifa

Female, 46 y.o., 04/05/1974
MRN: 18362946
Arrival date: 06/10/2020

Bed: 5B24

Consultant: Strife, Mandy

Treatment Team: Strife, Mandy; Chen, Liyi

Allergies: No known allergies

Patient Class: Outpatient

Current weight: 74kg

BMI: 28.5

Blood Type: A-positive

Informatics Consult Report:

Informatics Consult service: Who we are and what we do? (link)

Results are provided for patients with atrial fibrillation and cirrhosis. We considered warfarin treatment in these patients and outcomes of ischaemic stroke and mortality over 10 years.

Prevalence:

Number of individuals in cohort:

Prevalence of individuals having both conditions in England:

10 in 10,000 individuals

35,000 individuals scaled up to England's population

17.3% of cirrhotic patients have atrial fibrillation

1.3% of patients with atrial fibrillation have cirrhosis

Prognosis (1-year mortality):

Cirrhosis: 6%

Atrial fibrillation: 4%

Cirrhosis and atrial fibrillation: 15%

Queue patients for randomised trials:

Yes

No

Summary of evidence on efficacy and safety:

Overall summary:

Safety: We found lower all-cause mortality among warfarin users

Efficacy: We found lower stroke risk among warfarin users

Randomised trials:

Reported: 0

Recruiting: 0 (4 in patients without cirrhosis)

Meta-analysis of observational evidence:

Prior observational studies: 4

Warfarin use is associated with 29% reduction in stroke (non-significant)

Target trial emulation results: [See other outcomes](#)

Treatment variation - Warfarin initiators: 43%

All-cause mortality: 39% reduction

Ischaemic stroke risk: 73% reduction

Genetic evidence (Mendelian randomisation):

No studies on patients with atrial fibrillation and cirrhosis.

Evidence on drug target validation: warfarin use is linked to reduction of risk of large artery atherosclerotic stroke.

For full report, click here (link)

Chart Review

Results Review

Medications

Immunisations

Diagnoses

Orders

Results

Informatics Consult Request

Informatics Consult Report

Repeat consult

Want to know more? Schedule an appointment with the team.

Book appointment

Figure 3. Informatics Consult Report: Detailed trial evidence and currently recruiting trials of anticoagulation in patients with atrial fibrillation and cirrhosis to reduce stroke risk. (A) Clinical question and summary of trial evidence. (B) Previously completed and currently recruiting randomised trials evaluating anticoagulants and stroke outcomes have exclusion criteria related to cirrhosis.

A

Clinical question:
 Is oral anti-coagulation therapy safe and effective in reducing stroke risk in people with cirrhosis and atrial fibrillation?

Are there existing randomised controlled trials (RCTs) that provide sufficient evidence?

- Previously reported randomised trials on anticoagulants available for patients with cirrhosis = 0
- Currently recruiting anti-coagulant trials with stroke as an outcome = 4
- Currently recruiting trials available for patients with cirrhosis = 0

B

National Clinical Trial number	Trial acronym	Drug comparisons	Inclusion	Exclusion	Primary outcome
Previously completed trials					
NCT00262600	RE-LY	Dabigatran vs warfarin	Non-valvular atrial fibrillation	Patients with cirrhosis excluded	Stroke or systemic embolic event
NCT00403767	ROCKET AF	Rivaroxaban vs warfarin	Atrial fibrillation, history of a prior stroke	Patients with cirrhosis excluded	Stroke or systemic embolic event
NCT00412984	ARISTOTLE	Apixaban vs warfarin	Atrial fibrillation	Patients with cirrhosis excluded	Stroke or systemic embolic event
NCT00781391	ENGAGE AF-TIMI	Edoxaban vs warfarin	Atrial fibrillation	Patients with cirrhosis excluded	Stroke or systemic embolic event
Currently recruiting trials					
NCT03148457	ELAN	Early treatment vs late treatment with rivaroxaban, dabigatran, apixaban or edoxaban	Persistent atrial fibrillation and ischaemic stroke	Patients with cirrhosis excluded	Recurrent stroke, major bleeding, systemic embolic event
NCT03759938	OPTIMAS	Early treatment vs standard treatment with rivaroxaban, dabigatran, apixaban or edoxaban	Atrial fibrillation and acute stroke	Patients with cirrhosis excluded	Stroke, intracranial haemorrhage and systemic embolic event
NCT02618577	NOAH	Edoxaban vs aspirin	Atrial High Rate Episodes	Patients with cirrhosis excluded	Stroke, systemic embolic event, cardiovascular death
NCT01938248	ARTESiA	Apixaban vs aspirin	Sub-clinical atrial fibrillation	Patients with cirrhosis excluded	Stroke or systemic embolic event

Figure 4. Informatics Consult Report: New synthesis of prior observational evidence. Meta-analysis of the association between warfarin use and the risk of ischaemic stroke in observational studies including approaches for automation. (A) Characteristics of observational studies included in the meta-analysis. (B) Forest plot depicting the hazard ratios calculated with the DerSimonian and Laird random-effects models. HR = hazard ratio; CI = confidence interval; SE = standard error.

A

Study	Setting	Eligibility criteria	Number of patients with atrial fibrillation and cirrhosis	Warfarin use	No warfarin use	Adjusted estimates	Trial emulation
Kuo et al (2017)	Population	CHA2DS2-VASc score ≥ 2	9,056	754	8,302	Yes (propensity score)	No
Serper et al (2020)	Population	Patients without prior venous thromboembolic events	1,694	614	1080	Yes (propensity score)	No
Lee et al (2015)	Hospital-based	Patients with non-valvular atrial fibrillation, no previous diagnosis of mitral stenosis or prosthetic heart valve	321	173	148	Yes (multivariate Cox regression)	No
Choi et al (2017)	Hospital-based	Patients with non-valvular atrial fibrillation	465	113	352	No	No

B

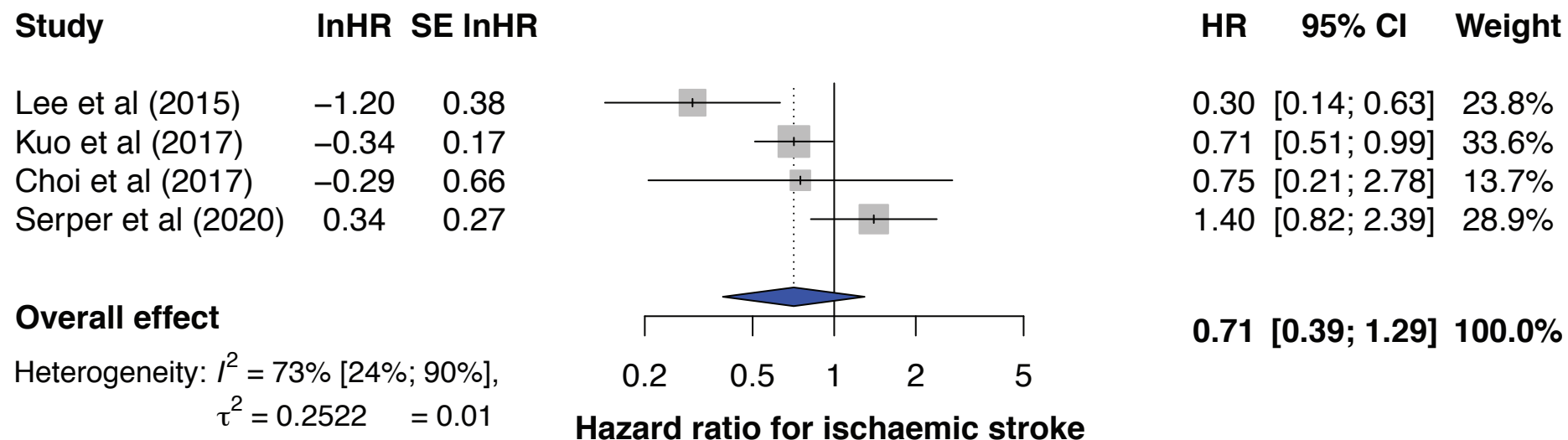


Figure 5. Informatics Consult report: New observational evidence through target trial emulation (intention-to-treat analysis) where eligibility and treatment assignment were aligned with time zero of follow-up, as is done in randomised controlled trials. (A) CONSORT diagram showing the selection of eligible individuals for the target trial emulation of anticoagulation therapy in patients with atrial fibrillation and cirrhosis. (B) Kaplan-Meier plots of the propensity-matched cohort for all-cause mortality and ischaemic stroke. Flow diagram depicts analysis design. P values from logrank tests were indicated. Hazard ratios from Cox proportional hazards regression analyses were indicated. Numbers in parentheses indicate the 95% confidence intervals.

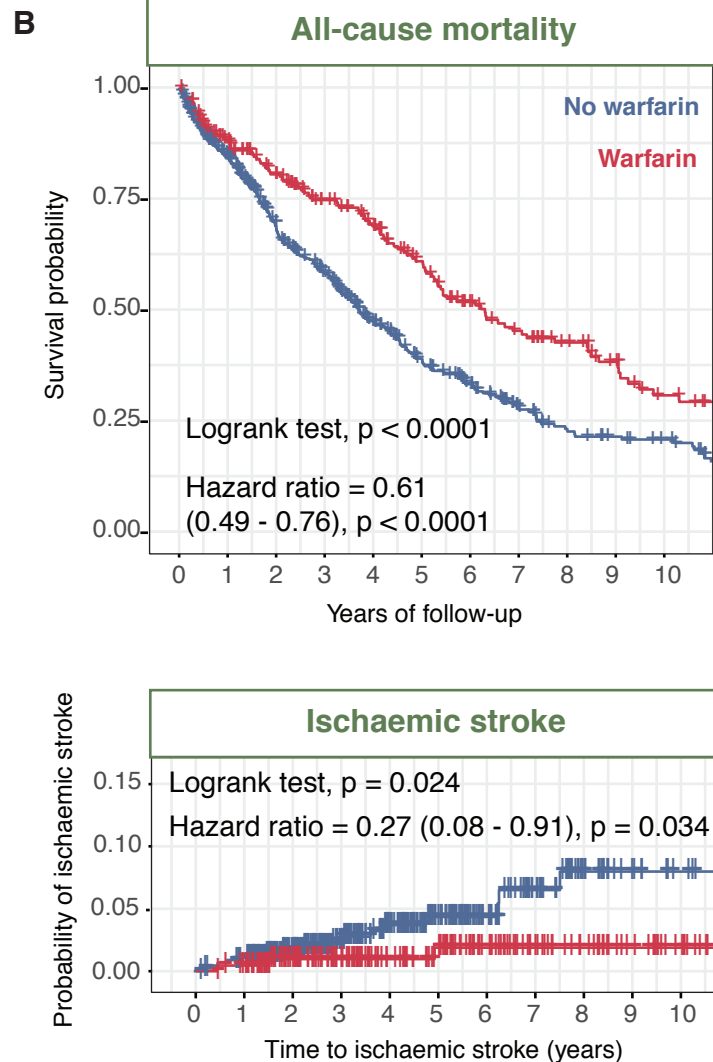
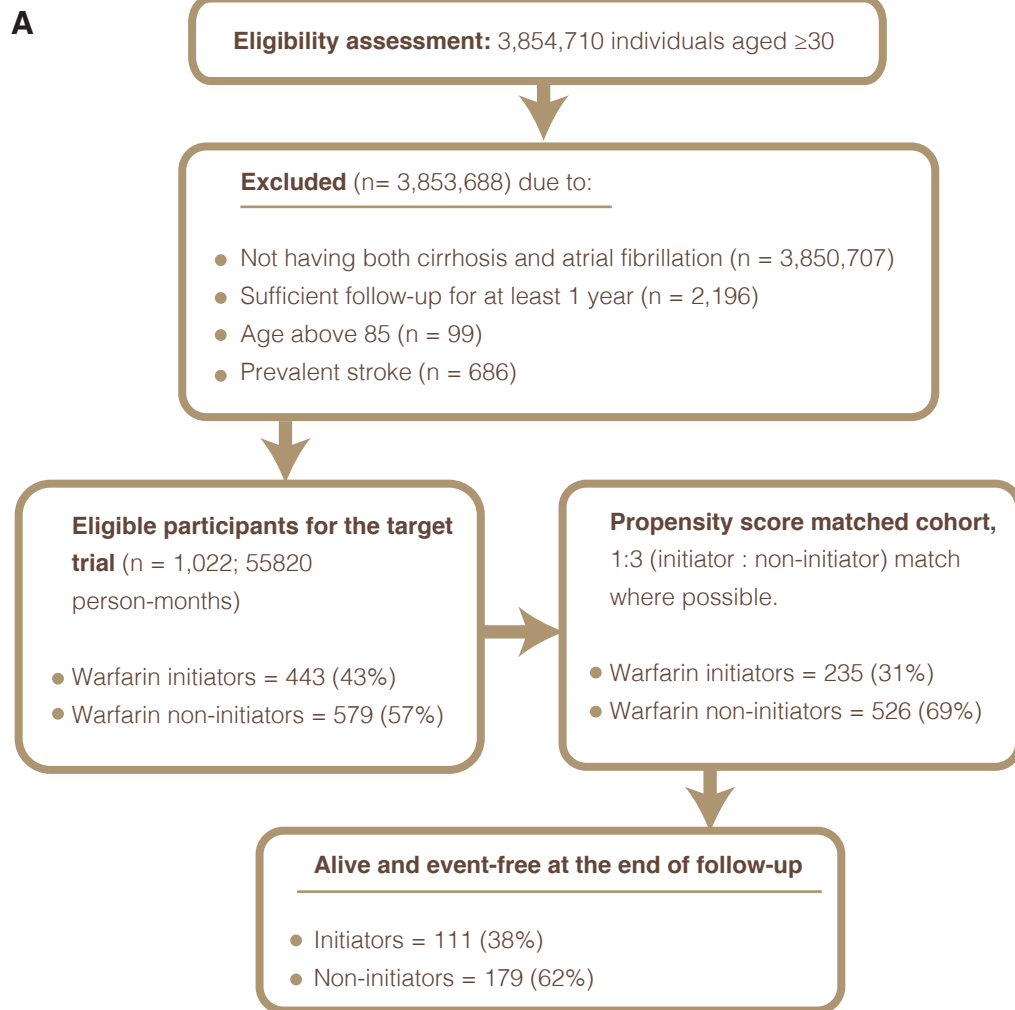


Figure 6. Informatics Consult report: Detailed genetic evidence.
 Two-sample Mendelian randomisation on circulating vitamin K1 levels and risk of stroke. * indicates significant results.

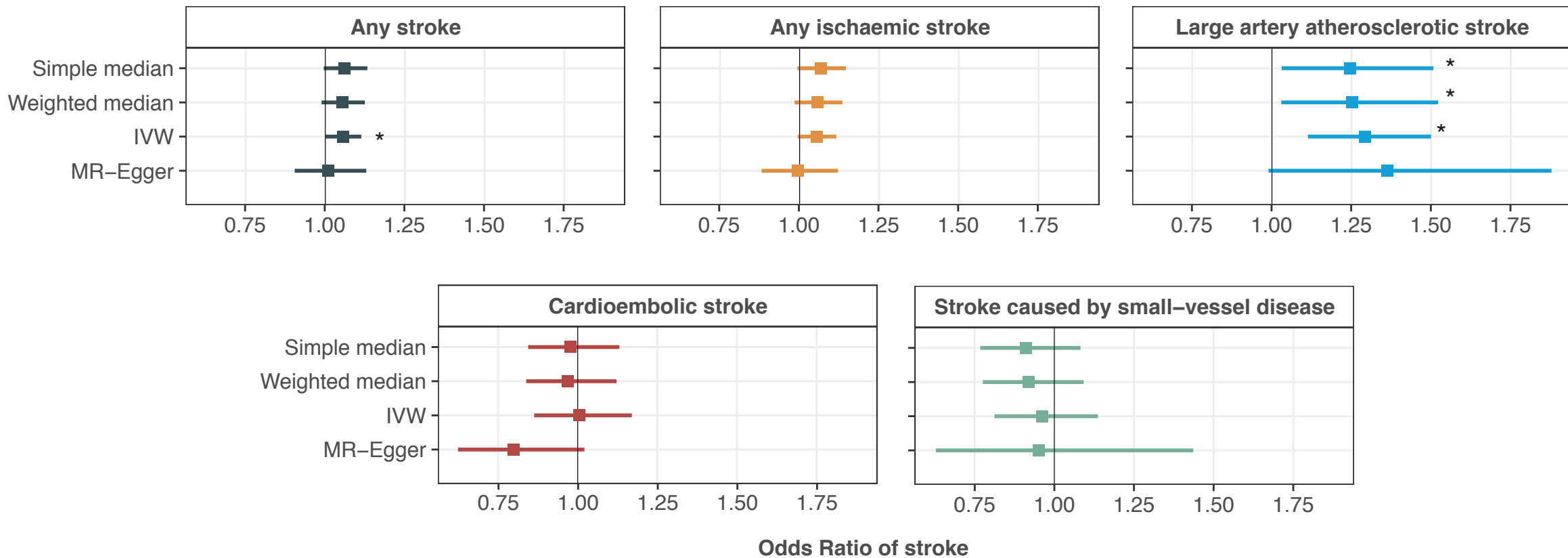


Table 1: Traditional versus automated approaches for evidence synthesis.

Task	Traditional approach	Approaches to automate
Evidence from randomised trials		
Search for recruiting trials	Perform search on the clinicaltrials.gov website. Requires manual decisions on relevant search terms.	Perform search on the clinicaltrials.gov website using search terms collated from free text input in the Informatics Consult platform. Potential to leverage developments on computable machine-readable trial protocols (https://doi.org/10.1177/009286150704100312) and computable phenotypes (i.e., algorithms to identify clinical characteristics derived from electronic health records) to identify potentially eligible patients for trial recruitment (https://doi.org/10.1161/CIRCOUTCOMES.119.006292). Potential to use sentence embedding and Google BERT as approaches for matching natural language queries with relevant trial protocols.
Summarise data of recruiting trials	Download search results from the clinicaltrials.gov website. Manually format tables. Extract additional information not present in downloaded data from the website by inputting NCT numbers.	Download search results from the clinicaltrials.gov website. Generate scripts for automated table formatting to retain relevant information. Create a Python web-scraping tool to extract free texts from specific clinical trials and return information on inclusion and exclusion criteria. Note that some websites do not allow web-scraping and exclusions may apply to the clinicaltrials.gov website.
Evidence from meta-analysis		
Search strategy	Requires manual decisions on relevant search terms.	Potential for mapping SNOMED-CT terms to MeSH descriptors used in PubMed (PMID: 17238584).
Identifying existing evidence from published sources and assessing eligibility.	Perform searches on PubMed. Manual curation and review of publications. Does not scale.	Semi-automated systematic reviews using machine learning and natural language processing for expedited evidence synthesis. For example, using 'bag of words' for classifying documents and using learned coefficients for predicting the probability of an unseen document. Examples of platforms for automating evidence synthesis include RobotReviewer and ExaCT, where the latter employs an information extraction engine that identifies and extracts text fragments that describe clinical trial characteristics on unseen articles (https://doi.org/10.1186/s13643-019-1074-9 ; https://doi.org/10.1186/1472-6947-10-56).
Extracting data and performing the meta-analysis.	Manual extraction of relevant tables and information. Not practical for batch extraction of data.	Semi-automated tool for converting PDF documents to XML using a rule-based system such as PDFX. Batch extraction of data from PDF documents can also be performed using the open-source CyberPDF, which improves the accuracy and efficiency in batch data processing. Extracted data is formatted into data frames for subsequent meta-analysis using the meta package in R or other existing packages. (https://doi.org/10.1145/2494266.2494271 ; https://doi.org/10.1145/3278576.3281274)
Evidence from target trial emulation		
Specifying the target trial protocol	This process requires a discussion between the clinician and informatician to determine the appropriate criteria, treatment strategies and outcomes.	Previous insights on specifying the target trial protocol can be collated automatically and be used to inform future target trial designs.
Cohort creation based on eligibility criteria in the target trial protocol	Manual cohort creation for each target trial. Does not scale.	This process can be pipelined using several functions to create cohorts in a consistent format with the covariates of interest. The DExtER tool for automated cohort creation can be employed.
Propensity score matching to match initiators and non-initiators	Once a cohort is created in the correct format containing all the covariates of interest, propensity score matching can be performed using the MatchIt package.	Additional approaches for causal inference analyses, including causal machine learning using the targeted maximum likelihood estimation approach can be investigated and pipelined.
Descriptive summary of the cohort before and after matching	The tableone package can be used to generate the baseline tables before and after propensity score matching.	Previous descriptive summaries on other related studies can be collated and featured in future target trials that investigate related clinical queries.
Cox regression on the matched cohort	Cox regression analyses is performed by fitting the coxph function using the survival package.	Additional regression analyses can be automated into the pipeline.
Kaplan Meier analysis on the matched cohort	Survival or cumulative incidence curves are plotted using the survminer package.	This can be pipelined to look at multiple outcomes at a time.
Scaling to other examples and datasets	Limited tractability.	Pipeline scalable to other datasets for cohort generation. Free text input from the Informatics Consult request form and report will inform additional opportunities to scale to other clinical questions.
Genetic evidence		
Identifying genetic variants associated with the exposure (e.g., drug) or risk factor in genome-wide association studies (GWAS). Identifying GWAS summary data for genetic variants associated with the exposure and outcome for a two sample Mendelian randomisation analysis.	Manual curation of GWAS summary data. Literature search for published genetic variants for the risk factor.	The MR-base platform for Mendelian randomisation can be employed to rapidly identify instruments for the exposure and outcome using GWAS summary data from their catalog. Additional GWAS summary data can be obtained from PhenoScanner, EMBL-EBI GWAS catalog and Integrative Epidemiology Unit OpenGWAS database.
Performing Mendelian randomisation.	Extract and format data identified above. Run Mendelian randomisation in R.	MR-base also includes an analytical platform for performing MR analysis. For exposures and outcomes not available in MR-base, this process can be pipelined to transform the GWAS summary data from other public sources into an analysis-ready format. Mendelian randomisation can be performed using the MendelianRandomisation package.

medRxiv preprint doi: <https://doi.org/10.1101/2021.01.10.21249331>; this version posted January 15, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Figure S2. Summary of survey responses.

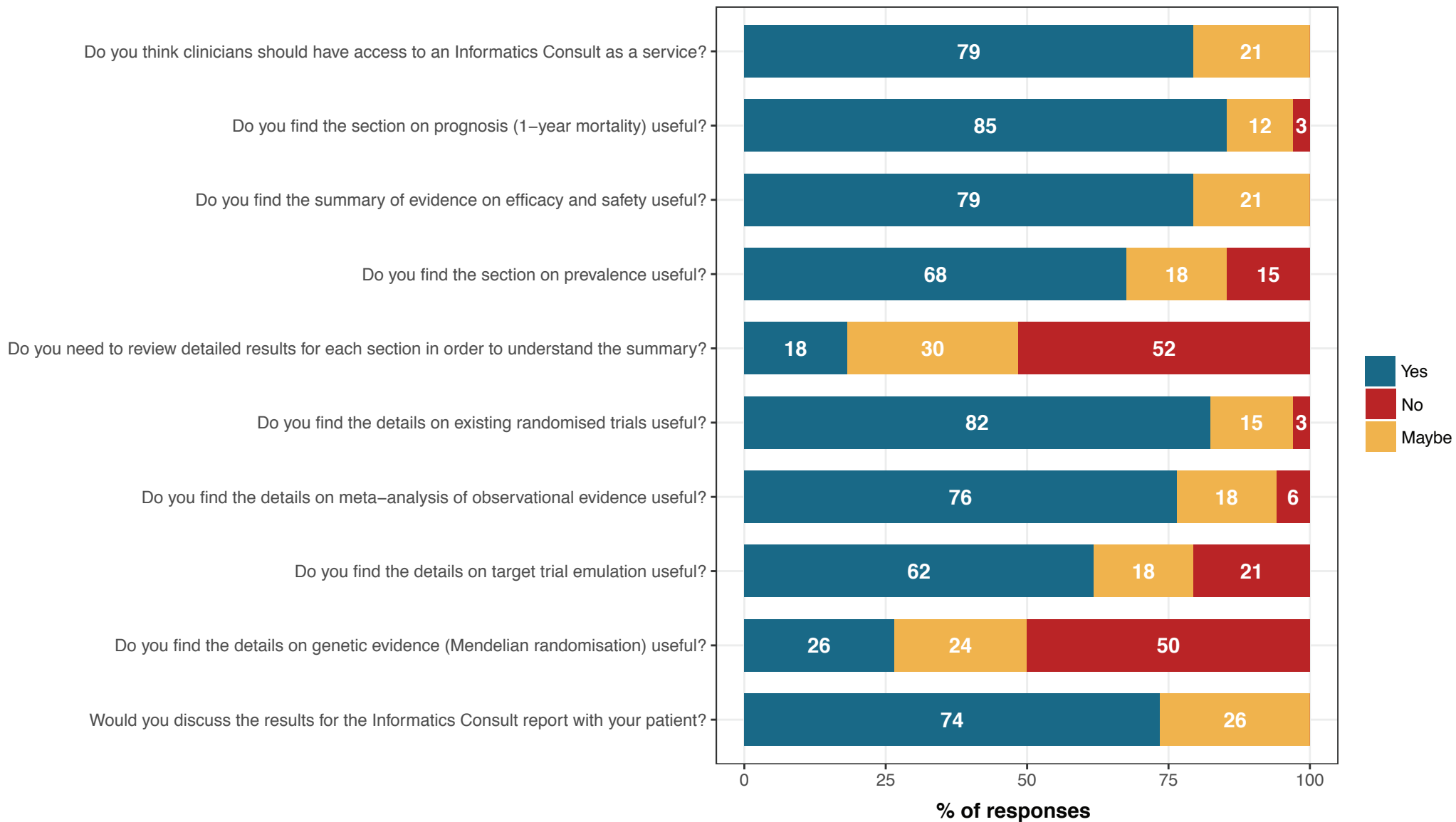


Table S1: Comparison of the Informatics Consult with traditional approaches to evidence generation.

Areas	Traditional methods of evidence generation	Informatics Consult
Who initiates the question?	Top down: typically driven by a small number of researchers	Bottom up: democratic, any clinician, any uncertainty and potentially patient driven
Simultaneous delivery of all relevant evidence	Not often possible, one study and one design at a time	Yes
Timescale for generating evidence	Months or years from question to answer	Hours, within clinical timescale
Embedded in health systems (the question or the answer or the means of providing answer)	No	Yes
Clinical trial prioritisation, informing trial design and recruitment	Top down: driven by specific groups of individuals	Bottom up: opportunities for embedding within clinical practice
Involvement of individuals	Researcher centric	Patient and clinician centric
Strength of evidence and levels of evidence	Levels of evidence classified in all major clinical guidelines; however, these are all based on peer reviewed published evidence	Same levels of evidence apply; however, the framework aims to return evidence within clinical timescales to mitigate hurdles in the journal peer-review process. Potential for collating information from past Consults in an open-access repository.
Clinical guideline recommendations based on evidence	Yes, information from randomised trials and peer-reviewed publications are included in clinical guidelines	Raises important questions on what basis might clinicians follow the Consult before it is peer reviewed and published
Regulatory approvals on the labels of drugs and medical devices	Food and Drug Administration (FDA), European Medicines Agency and Medicines and Healthcare products Regulatory Agency established processes for trials and for the Real-world Evidence framework	Largely untested