How could equator-network.org be improved? Inferences from Google Analytics data.

Commit ID: 16a22c3

2023-08-01

|  |
| --- |
| This is the same as the article I drafted a year or so ago. Gary had suggestions about how to reframe it but I’ve not edited it at all yet. |

## Abstract

The UK EQUATOR Centre maintains a website to help authors find and access reporting guidelines. Web analytics shine a light on how well the website is performing and what could be done to improve it. Here we describe some key insights from web analytics. 1) Many users leave the site very quickly without interacting with it at all. 2) Guideline views are very heavily skewed towards the few guidelines that are featured on EQUATOR’s home page and on journal instructions to authors. 3) Users access checklists more often than the full guidance, perhaps because the full guidance is harder to find. 4) EQUATOR could optimize their site for search engines. 5) EQUATOR could improve acquisition data by updating their links on third party websites and offline documents. 6) Some countries with high publication volumes are under-represented amongst website visitors. 7) EQUATOR could consider translating parts of their website to cater for a global audience.

## Introduction

In chapter **?var:chapters.survey\_content** I found two studies that exploring authors’ awareness of EQUATOR as an organisation ([1]; [2]). Although some authors will know EQUATOR from their training programmes or publications, most will know EQUATOR from its website. Despite running on a shoestring budget and without any in-house expertise in software, design, or user experience, the website’s traffic has increased from 100,000 users to almost 1 million over the last 10 years, attracting visitors from all around the world. For context, around 4.3 million authors published 2.4 million academic articles in 2014 [3], and the STM association estimated that 11 million people working in research and development in 2018 [4]. Whilst these numbers will have grown since, they suggest that EQUATOR’s website traffic numbers are within a single order of magnitude of its potential audience size.

Hence what started as a research-group’s attempt to catalogue reporting guidelines has become a significant part of the academic eco-system. And yet, EQUATOR has never formally considered how successfully their website is performing, or how it could be improved.

In this chapter I describe how I worked with EQUATOR to identify key metrics of success and how I used web analytics to describe those metrics. In the discussion section, I infer how the website may need improving.

## Methods

Google Analytics [5] is a web analytics service that helps website owners track and understand users’ activity. Used by 85% of websites globally, it is the most popular web analytics service by far [6]. EQUATOR has used Google Analytics to collect data since creating their website. Mostly they use the data to report high level impact metrics to funders (such as number of visitors) but they have never used it to evaluate their site in depth.

When evaluating a website, there are a huge number of metrics you can consider; Google Analytics collects over 50 by default [7]; [8], and Google Tag Manager [**WebMobileTag?**] allows you to collect additional custom metrics (see **?@tbl-ga** for definitions of Google Analytics terms and metrics). With so many options, the first step of evaluation is to decide what metrics are most important for your website’s objectives.

I met with three members of the UK EQUATOR Centre in October 2021. When I asked what the main purposes of their website [9] was, I received many answers. EQUATOR has content aimed at authors, editors, peer reviewers, educators, librarians, and guideline developers. There are web pages promoting training courses, toolkits for writing and reviewing research, newsletters, and blogs highlighting work done by EQUATOR and guideline developers.

However, at its core, the EQUATOR staff I spoke to agreed that the purpose of the website is to help the global research community learn about and access reporting guidelines. They want website visitors to access the guidance that is right for them and come back to the website whenever they need guidance.

To explore how far reality meets this vision, I used Google Analytics to answer the following questions:

* How many people visit the website each year?
* Where (in the world) are visitors from?
* How often do visitors come back?
* How do visitors get to the website?
* How many visitors access guidance?
* What guidelines do visitors access?
* How many visitors access publications or checklists?

The last two questions could not be answered by Google Analytics’ default configuration; it recorded which database records visitors looked at, but not how many people went on to view guidance on third party websites, or whether visitors were viewing checklists or full guidance. Therefore I used Google Tag Manager to create two custom metrics: one which counted when visitors downloaded a reporting checklist file, and another which counted when visitors accessed a third party website.

EQUATOR staff felt it was important that website users were able to access the *right* resource. Although Google Analytics could tell us *what* pages visitors access, it can’t tell us what they *needed*, or *why*. To answer this, I used PopupSmart [**PopupBuilderThat?**] to create a single-question exit survey that asked “What were you looking for today?”. I used a single, open ended question to maximize my response rate.

## Results

## Keep visitors engaged

Google Analytics classified almost all users as new, which means they had not visited the site within the previous 2 years (the default expiry limit for Google Analytics cookies). Most users only visited the site once within 2021. Two thirds of sessions lasted less than 10 seconds, and over half of sessions ended without the user interacting with the site at all. EQUATOR’s website holds users’ attention for a very short amount of time and, currently, a large proportion of users are leaving the site without accessing any reporting advice.

A third of sessions begin with users arriving on the home page, where almost half of visitors then immediately leave without interacting with the website. Google Analytics calls this behaviour bouncing. Many users arrive directly on a guideline database page. Some of these pages have even higher bounce rates, such as 56% for STROBE and 73% for COREQ.

There are many reasons why a user may abandon a website quickly. In autumn 2021, the UK EQUATOR Centre added a popup survey to ask users why they were leaving the website. Whilst some answered that they “got what they were looking for,” others wrote that “the site is very complex,” “too big a mess” or that they “could not find what they needed.” One visitor wrote that they were looking for the formatting instructions for their target journal, suggesting that they did not understand what EQUATOR was or that they were no longer on the journal’s website. EQUATOR should continue to research why users leave these pages so quickly, why they don’t return and how the website could be improved, with special attention paid to the home page and guideline pages.

## Help users discover reporting guidelines besides the few that are featured

EQUATOR maintains a database of reporting guidelines. Each database entry is viewable as a webpage that displays the name of the guideline, bibliographic data, meta data, and links to the associated publications and files. Users must navigate this page before accessing the guideline or checklist.

There are over 500 reporting guidelines indexed in EQUATOR’s database, but hardly any are viewed regularly. Only 13 guideline database records were viewed in more than 1% of visits, all of which appear in the list of “reporting guidelines for main study types” featured on EQUATOR’s home page and in a side bar on all reporting guideline sub-pages. STROBE was the most viewed reporting guideline, with 231,207 unique page views in 2021. These numbers dropped rapidly: SQUIRE, the tenth most viewed reporting guideline record, was viewed ten times less frequently than STROBE. Only 65 reporting guideline pages were viewed in more than 0.1% of sessions.

Viewership is thus heavily skewed towards the few guidelines that are featured on EQUATOR’s home page and journal instruction to author pages. We expect some guidelines to be viewed more than others, but this distribution seems extreme and suggests that authors are not discovering newer or more specific guidelines that might be more applicable to their work.

Possible solutions to this would be to make the search function more prominent and easier to use, and to better highlight related guidelines at the top guideline databases pages, checklists, and guidance articles, along with clear instructions of when each reporting guideline should or should not be used. Guideline developers should be aware that inclusion in EQUATOR’s database does not guarantee authors will find and use their resources, and so should consider complementary dissemination strategies.

## Do users access checklists, articles, or neither?

Checklists are accessed more frequently than the full guidance when they are available (Table 3). This ratio is 15:1 for STROBE, 6:1 for CONSORT, and 8:1 for STARD. This is probably because EQUATOR links to checklists at the very top of the guideline page and links to the full guidance appear lower in the page and are harder to find. This has implications for guideline developers who should ensure that important information about the guideline such as its aim, scope, or how to use it, should be placed in the checklist, as should a link to the full guidance. Guideline developers should not assume that users will discover the E&E document before the checklist or at all.

It is worth noting that many users leave the guideline reporting guideline database page without accessing any resources. Even if a visitor clicks an outbound link to the full guidance, these links generally take visitors to a PubMed record where they must press another button to access the full text. There will therefore be additional drop-off before visitors reach the full guidance. EQUATOR should improve their reporting guideline database pages to make the checklists and full guidance easier to discover.

## Search engine optimization to catch naïve authors early in their research

A third of users arrive at the site via a search engine. There are many ways EQUATOR could optimize their site for search engines, including adding metadata, mobile optimization, and taking advantage of Google Search’s featured snippets and description features. Keyword rankings could also help to drive potential users to the site. For instance, Google Search Console is a Google product that allows website owners to view how their site performs in Google searches. It shows that when users search for “STROBE guidelines,” EQUATOR’s site appears at the top of search results and has a click through rate of 36%. However, if a user searches for “how to write an epidemiological report,” EQUATOR drops to 29th place with a click through rate of 0%. EQUATOR should ensure the site is optimized for naïve users at an early stage of writing who may not know guideline acronyms.

## Update links on other websites to understand where users are coming from

An eighth of traffic was explicitly labelled as referrals from other websites, most commonly Wiley and Elsevier journals and Manuscript Central. However, it is likely that much of the traffic from unknown source is also referrals from other websites. These unknown sources currently account for half of EQUATOR’s traffic. Some of this unknown traffic may come from links within offline files. It is likely that a lot is from journal websites and submission systems that link to EQUATOR using links that start with *http* instead of *https*. When a secure website with an *https* address links to a less secure site with an *http* address, no referral data gets sent. EQUATOR upgraded its website to use *https* years ago, but other websites have continued to link using *http*. A campaign to have journals and submission platforms use *https* links would result in more correct referral data. This would tell EQUATOR which journals are successfully recommending reporting guidelines and would allow EQUATOR to infer visitors’ intentions. For example, traffic from submission systems may signify users who have been asked to complete a checklist from.

## Target authors from under-represented countries

Over 800,000 users visited the EQUATOR website in 2021 (Table 4), 150,000 more than the previous year, thus continuing growth seen since attracting 20,000 visitors in 2008. A third of users were from the United States or United Kingdom. Brazil accounted for the same proportion of users as the UK (7%) despite producing far fewer citable, medical documents [10]. Conversely, China produces twice as many citable documents as the UK but accounts for only 5% of users. These numbers suggest that awareness of EQUATOR is greater in some counties than in others. EQUATOR should continue to run awareness campaigns and should prioritize countries like China which have high publication output but low visitor numbers.

## Consider non-native English speakers

Two fifths of users had their browsers set to a language other than English (Table 4), most frequently Portuguese, Chinese, and Spanish. Manually translating popular guidelines or website sections may be appropriate for frequent languages, but manual translations are expensive, difficult to update, and cannot be scaled. Automatic machine translation is more scalable, easy to update, and inaccuracies can be refined with custom glossaries and language models. EQUATOR and guideline developers should also use plain language to help users and automatic translators alike.

### Survey data suggest some visitors may not understand what the website is about, and may not find it useful

I let the survey run for 2 weeks before deciding to take it down.

The response rate was poor -

The responses that we received were rarely insightful. Only three visitors answered the question. One wanted “to get reporting guidelines” another wanted “a tool called standard for reporting qualitative research”, and a third “could not find any reporting guidelines for reporting guidelines (specifically, abstracts for reporting guidelines)”. Two others hinted at what they wanted but weren’t explicit, “qualitative”, or “word format would be more easy to fulfil” (presumably referring to reporting checklists). A handful of users didn’t specify what they were looking for, but stated whether they had found it (e.g. “got what I was looking for!”, “found what I needed!”, “Could not find what I needed”, “I cannot find the guidance that I seek”, “I don’t see what I want”).

But the majority of responses did not help answer our question. Some visitors seemed to be in the wrong place: one was looking for a “quality of life questionnaire”, another for “scientific research”. Two authors seemed to be looking for requirements for specific journals: one wrote “format for paper submission to Hindawi”, and another wrote “awful site. I just want to know the requirements in terms of number of words and format for a submission and cannot seem to find this anywhere”. Other visitors also seemed frustrated with the website (“i did not under stand any thing”, “The site is very complex”, “Too big a mess”) and the popup itself (“You might want to do something about your annoying popup!”, “don’t ask that”).

Some user’s responses made no sense at all. One wrote “ALGERIA”, another “Germany”.

## Limitations

Google Analytics uses cookies to track users over time. If a user clears their cookies between visits or uses multiple devices or browsers, the user will appear as multiple users. Cookies expire after 2 years by default. The proportion of new vs. returning users is thus an overestimation but, nevertheless, is still high.

Although we discuss bounce rates here it is difficult to know what a “good” bounce rate would be. E-Commerce websites may hope for a bounce rate of around 40%, but newspapers may consider 70% to be good. What is important is that EQUATOR now has a baseline against which to measure improvement. Ultimately, numbers can only tell you so much. Counting bounces is useful, but only qualitative research will explain *why* users bounce from the EQUATOR website with the frequency that they do.

## TODO Definitions

User

Visitor

Session

Bounce

Bounce Rate

STROBE

PRISMA

CONSORT

STARD

SQUIRE

E&E

Google Analytics

Google Search Console

### Table 1 Session information for EQUATOR-Network.org for the year 2021. A session is defined as group of user interactions that take place within a given time frame. If a user visited the site twice within a year, then that would appear as two sessions. A session ends after 30 minutes of inactivity.

|  | Number | Percent |
| --- | --- | --- |
| Number of sessions | 1,209,420 |  |
| Mean number of sessions per user | 1.5 |  |
| Number of sessions originating from… |  |  |
| a referral from another website | 142,158 | 12% |
| a search engine | 417,671 | 35% |
| unknown sources | 590,659 | 49% |
| Number of sessions lasting… |  |  |
| < 10s | 760,967 | 63% |
| < 1 min | 908,516 | 75% |
| Number of pages viewed within a session |  |  |
| 0 | 1,451 | 0% |
| 1 | 676,670 | 56% |
| 2 | 223,864 | 19% |
| 3 | 81,878 | 7% |
| 4 or more | 225,557 | 19% |
| Number of sessions ending without interaction | 648,370 | 54% |
| Number of sessions including views of… |  |  |
| reporting guidelines | 605,570 | 50% |
| landing page | 411,429 | 34% |
| library | 83,238 | 7% |
| toolkits | 15,574 | 1% |
| study design info | 15,268 | 1% |
| (All other categories of content were viewed in less than 1% of sessions) |  |  |

### Table 2 Pages where visitors began their session and their bounce rates between 1st of January and 1st of July 2022. Bounces are sessions where a visitor leaves without interacting with the website at all.

| Page | Sessions | % Sessions | Bounce Rate |
| --- | --- | --- | --- |
| All pages | 602,921 | 100% | 53% |
| Home page | 218,260 | 36% | 45% |
| STROBE | 68,274 | 11% | 56% |
| PRISMA | 29,398 | 5% | 60% |
| Reporting guidelines | 27,122 | 4% | 4**0%** |
| CONSORT | 24,615 | 4% | 55**%** |
| COREQ | 19,962 | 3% | 73**%** |
| STARD | 19,316 | 3% | 57**%** |
| SRQR | 12,039 | 2% | 65**%** |
| CARE | 10,759 | 2% | 56**%** |
| TRIPOD | 10,342 | 2% | 58**%** |

### Table 3 How frequently reporting guideline pages were viewed and resources accessed between the 1st of January 2022 – 1st of July 2022. Outbound links were mainly links to publications and files were mainly checklists, but a few guidelines also have flow diagrams which were rarely downloaded. The ten most accessed guidelines are shown.

| Reporting Guideline | Sessions where database page was viewed |  | Users that accessed an outbound link | Users that accessed a checklist |
| --- | --- | --- | --- | --- |
| STROBE |  | 110,910 | 2,502 | 38,780 |
| PRISMA |  | 50,146 | 17,240 | 15,522 |
| CONSORT |  | 45,832 | 2,465 | 13,862 |
| COREQ |  | 30,487 | 11,152 | - |
| STARD |  | 27,301 | 987 | 7,803 |
| SRQR |  | 26,132 | 10,035 | - |
| CARE |  | 21,892 | 8,761 | 8,350 |
| TRIPOD |  | 15,240 | 4,678 | 3,960 |
| SPIRIT |  | 11,108 | 3,985 | 3,244 |
| SQUIRE |  | 11,061 | 1,022 | 3,061 |

### Table 4 Google Analytics data for equator-network.org for the year 2021.

|  | Number | Percent |
| --- | --- | --- |
| Number of users | 830,134 |  |
| Number of users who had not visited before | 823,087 | 99% |
| Country (top 10) |  |  |
| United States | 195,217 | 24% |
| United Kingdom | 61,292 | 7% |
| Brazil | 55,894 | 7% |
| China | 41,628 | 5% |
| India | 41,196 | 5% |
| Australia | 30,647 | 4% |
| Canada | 29,620 | 4% |
| Netherlands | 25,426 | 3% |
| Germany | 24,561 | 3% |
| Spain | 23,020 | 3% |
| Language (top 10) |  |  |
| English (United States) | 371,331 | 45% |
| English (Great Britain) | 86,580 | 10% |
| Portuguese (Brazil) | 50,172 | 6% |
| Chinese | 42,951 | 5% |
| Spanish | 36,204 | 4% |
| French | 14,708 | 2% |
| Dutch | 14,331 | 2% |
| Spanish (Latin America and Caribbean) | 14,259 | 2% |
| Japanese | 13,381 | 2% |
| German | 12,460 | 2% |
| Users whose browsers are set to English\* | 498,286 | 60% |
| Number of sessions that users made within the time window |  |  |
| 1 or more sessions | 818,512 | 99% |
| 2 or more sessions | 176,570 | 21% |
| > 3 sessions | 72,886 | 9% |

\* Includes 105 English locales like United States, India, Ireland etc.

# References

{"Analytics Tools & Solutions for Your Business – Google Analytics", n.d.}

{"Scimago Journal & Country Rank", n.d.}

{"The EQUATOR Network | Enhancing the QUAlity and Transparency Of Health Research", n.d.}

{"The EQUATOR Network Google Analytics", n.d.}

1. Fuller T, Pearson M, Peters J, Anderson R (2015) [What affects authors’ and editors’ use of reporting guidelines? Findings from an online survey and qualitative interviews.](https://doi.org/10.1371/journal.pone.0121585) PLoS ONE 10:e0121585

2. Gi̇ray E, Coskun OK, Karacaatli M, Gunduz OH, Yagci İ (2020) [Assessment of the knowledge and awareness of a sample of young researcher physicians on reporting guidelines and the EQUATOR network: A single center cross-sectional study](https://doi.org/10.5472/marumj.682337). Marmara Medical Journal 33:1–6

3. Plume A, Weijen D van (2014) Publish or perish? The rise of the fractional author…. Research Trends 1:

4. STM Association STM Global Brief 2021 Economics & Market Size.

5. Analytics Tools & Solutions for Your Business Google Analytics. Google Marketing Platform

6. Usage Statistics and Market Share of Traffic Analysis Tools for Websites, July 2023.

7. [GA4] Automatically collected events - Firebase Help.

8. [GA4] Predefined user dimensions - Firebase Help.

9. The EQUATOR Network | Enhancing the QUAlity and Transparency Of Health Research.

10. Scimago Journal & Country Rank.