ESGEM-AMR Working Group: Interpretive Standards for AMR Genotypes

Overview

ESGEM-AMR is a working group coordinated by the ESCMID Study Group for Epidemiological Markers (ESGEM). It is led by Prof Kat Holt (LSHTM) and Dr Natacha Couto (ESGEM Chair), and is open to ESGEM members and others with relevant expertise.

Purpose

The overall purpose of the group is to **capture expert knowledge** on the relationship between antimicrobial resistance (AMR) genotypes and antimicrobial susceptibility testing (AST) phenotypes in bacterial pathogens, in a manner that:

i. Recognises and accounts for differences between species (is organism-specific);
ii. Connects with EUCAST Expert Rules, Expected Phenotypes, and other standards as far as practicable;
iii. Uses standardised data structures to capture expert knowledge, which are interoperable with a range of informatics tools and databases (i.e. not platform-specific).

The initial focus of the ESGEM-AMR working group will be expert curation of interpretive standards for AMR genotypes in the AMRrules format, but may expand in future to include other activities/projects. The group will seek to engage with EUCAST and other relevant groups within ESCMID and externally.

Membership

Anyone with relevant expertise can apply to join the ESGEM-AMR working group. Prospective members will be asked to register their interest via an online form, to be reviewed by the ESGEM-AMR leads and ESGEM Executive Committee, and will be selected on the basis of expertise and ensuring a spread across organisms.

Working group members will be required to sign a Memorandum of Understanding outlining the roles and responsibilities of the working, and a Code of Conduct, to formalize their involvement with the working group.

Active working group members will be encouraged to note their membership on their CV, and will be invited to co-author outputs.

As a working group of ESGEM (the ESCMID Study Group for Epidemiological Markers), ESGEM-AMR members will be encouraged to also join ESGEM, which requires current membership of ESCMID, however this is not a requirement.

Scope

The initial focus will be on clearly delineating **'wildtype' (core) genotypes underlying 'wildtype' (intrinsic/expected) phenotypes** for clinically relevant bacteria. This is analogous to EUCAST's Expected Resistance and Expert Rules for susceptibility testing, which capture expert knowledge on interpretive rules for AST, and ideally will capture genetic mechanisms behind all expected resistances.

Subsequently, we envisage addressing rules concerning the interpretation of **acquired determinants**, including genes, mutations and **combinations** thereof, and how these should be interpreted in terms of expected **deviation from wildtype phenotypes** in each organism.

The scope of organisms to be included will depend on the availability of expertise to contribute to the working group, but we will aim to first develop rule sets for:

1. ESKAPEE pathogens:
   Enterococcus faecium
   Staphylococcus aureus
   Klebsiella pneumoniae species complex*
   Acinetobacter baumannii
   Pseudomonas aeruginosa
   Enterobacter cloacae complex*
   Escherichia coli

   (*Note the ESKAPE list was established prior to WGS-informed understanding of the structure and diversity of the *Klebsiella pneumoniae* or *Enterobacter cloacae* species complexes, so referred to *Klebsiella pneumoniae* and *Enterobacter* spp.)

2. Other organisms on the WHO Priority Pathogens list:
   Salmonella spp., *Shigella* spp., and other *Enterobacteriaceae*
   Neisseria gonorrhoeae
   Streptococcus pneumoniae
   Haemophilus influenzae
   Helicobacter pylori
   Campylobacter spp.

   (Note the current list dates from 2017 and an update is due Q2 2024, priorities will be reviewed once it is released)

3. Other organisms of clinical relevance where sufficient expertise and data is available, prioritizing those with Expected Resistant phenotypes

Scope of Work

Individual members or subgroups will be **allocated to work on specific organism/s** that match their expertise. The main task will be to **propose rule sets** for the assigned organism/s, in the AMRrules format (see Table 1 below and Technical Guidance). Preparation of publications describing the rationale and testing for a given organism or set thereof will also be encouraged.

Ideally, the interpretive standards should capture all exceptions to the generalized interpretation of 'presence of gene X' implies nonwildtype resistant (R^{NWT}) to drug Y'. Ultimately, they should also differentiate R^{NWT} from I^{NWT}, and interpret combinations of genes. However, the initial priority will be creating rule sets that clearly delineate **core genes** associated with **'wildtype' phenotypes** for each species (see AMRrules overview for the reasons for this approach).

Working group members will be encouraged to identify public matched genome/AST data to use in validation and testing, and to seek or share additional unpublished data if possible. If unpublished data is shared for this purpose, it must be treated in confidence under the terms of the working group's Memorandum of Understanding and Code of Conduct, which all members must sign before starting. The publishing and public deposition of data is strongly encouraged, either independently (so it can be cited by the working group publication) or as part of the working group (if this suits the needs and timelines of the data owners). However using or sharing unpublished data for the purpose of setting or testing rules by the working group does not obligate anyone to publicly release data, this will always be the decision of the data owners.

Once a proposed rule set for a given organism has been drafted, it should be submitted for review by the ESGEM-AMR chairs and lead bioinformatician (Jane Hawkey, Monash), according to agreed criteria. Approved rule sets will be added to the master scheme and included in the next release.

Target organisms are outlined above.

Technical details

See: ESGEM-AMR Technical Guidance

Work plan and timeline

See ESGEM-AMR Workplan and Timeline

Out of scope: Predictive classifiers

There is intense academic and commercial interest in training classifiers to predict AMR phenotypes from different kinds of genomic features (e.g. kmers), however the working group will focus on the specific issue of interpreting genotype profiles, generated by screening WGS data against a database of known AMR determinants as discussed above (e.g. screening against the refgenes AMR database using the AMRfinderplus tool). Genotype profiling is the most widely used and accessible approach to investigating AMR from WGS data, is based on decades of accumulated knowledge of molecular mechanisms of resistance, and the resulting profiles are valuable in understanding the mechanisms and spread of resistance in addition to their value in predicting phenotypes.

Planned Outputs

Rule sets

The key output of the ESGEM-AMR working group will be organism-specific rule sets with which to interpret AMR genotypes. We will adhere to FAIR principles to ensure the outputs are findable, interoperable, accessible, and reusable. Rule sets will follow a **standard format** (see Table 1 below and Technical Details). Details of the rule specification will be further developed and refined by the working group during the course of their activities. Rule sets will be made **publicly available** via open-access repositories under a permissive license (GNU General Public License v3.0). They will be **versioned** via numbered releases, and issued with stable document object identifiers (DOIs).

Table 1. Proposed format for organism-specific rules, populated with example rules for *Klebsiella pneumoniae*.

species	allele	context	drug	category	PMID	note
Klebsiella pneumoniae	blaSHV	core	ampicillin	wt R	32284385	Specific alleles can be ESBL, these are mostly mobile
Klebsiella pneumoniae	oqxA	core	ciprofloxacin	wt S	30834112	Wildtype core gene, not expected to confer multiple drug resistance unless mobilised under strong promoter
Klebsiella pneumoniae	oqxB	core	ciprofloxacin	wt S	30834112	Wildtype core gene, not expected to confer multiple drug resistance unless mobilised under strong promoter
Klebsiella pneumoniae	fosA5_fam	core	fosfomycin	wt S	27261267	Wildtype core gene, not expected to confer multiple drug resistance unless mobilised under strong promoter

Rule sets and associated tools should be **compatible with existing resources** for AMR genotype analysis as far as possible, including use of common sequence identifiers, standard gene nomenclature, and data formats. Preliminary code has been developed for annotating the gene-level reports output by the AMRFinderPlus tool, which uses the NCBI refgene database of AMR determinants. Working group outputs should be interoperable with the hAMRonization format, to facilitate compatibility with the outputs of CARD RGI, ResFinder, and >12 other AMR genotyping tools whose outputs can be readily converted to hAMRonization format.

Table 2. Example AMRFinderPlus genotype report for a *K. pneumoniae* (ERR257656), annotated with additional columns using the rule set shown in **Table 1**. Note a column reporting the source of the organism-specific interpretation is also annotated ("Klebsiella pneumoniae; v1.1") but is not shown here for brevity.

Gene symbol	Context	Org interpretation	Drug	Class	Subclass	% Coverage	% Identity
oqxB19	core	wt S	ciprofloxacin	PHENICOL/QUINOLONE	PHENICOL/QUINOLONE	100	100
oqxA3	core	wt S	ciprofloxacin	PHENICOL/QUINOLONE	PHENICOL/QUINOLONE	100	100
blaSHV-33	core	wt R	ampicillin	BETA-LACTAM	BETA-LACTAM	100	100
fosA	core	wt S	fosfomycin	FOSFOMYCIN	FOSFOMYCIN	100	100
dfrA1			trimethoprim	TRIMETHOPRIM	TRIMETHOPRIM	100	100

Publications

The overall motivations, goals, approach and rule sets will be described in an article co-authored by all contributors to the working group.

Additional articles describing the details for specific organisms or sets thereof, or describing associated tools or protocols, may also be developed by the working group, with authorship to be determined by contribution.

Wherever possible, publications should be accompanied by accessions for publicly available genome sequence and AST data supporting the rule development and testing. However the timing of this remains the decision of the data owners as outlined above, and in the Memorandum of Understanding.

Conferences

The working group will aim to present their activities at relevant conferences, upcoming ones that we plan to target include:

ASM-NGS (Oct 13-16, 2024, Washington DC; late-breaker abstracts Aug 14 2024)
ECCMID (April 12-15, 2025, Vienna; abstracts due 27 Nov 2024)
ABPHM (May 21-23, 2025, Hinxton; abstracts due ~Jan 2025)
IMMEM (Sep 17-20, 2025, Porto; abstracts due ~Mar 2025)

Work Plan and Timeline

April 2024 — Working group launch and call for interest - ESGEM session @ ECCMID

May 2024 — Introductory webinars

- 2x, to accommodate different time zones
- Advertised to members of ESGEM and other ESCMID study groups; SEDRIC, PHA4GE, microbiology societies
- Introduce project goals, mode of working, commitment needed, benefits
- Invite questions and feedback
- Invite attendees to complete online form to register EOI

2 June 2024 — EOI form due

Mid June — Review EOI submissions

- Identify submissions with suitable expertise
- Review spread of expertise and decide individual bugs / groups of bugs / whether to appoint / invite subgroup leads
- Invite selected submissions to join working group (by late June)
- Members sign MOU and Code of Ethics

July — Initial meeting of working group

- Review members, goals, modes of working
- Data formats, protocols, support tools
- Discuss standards and process for review of rules
- Set goals and timelines for next 6 months
- Establish comms model - slack? - especially to connect within subgroups, but also discuss issues around formats, coding, analysis, etc across bugs

September — Monthly progress meetings

- Issues emerging related to protocols, data standards, dictionary
- Updates on each organism
- Ideas for outputs, conferences, publications, funding

April 2025 (?) — Manuscript submission & presentation at ECCMID