

Drug Safety

Application of Information Retrieval Approaches to Case Classification in the Vaccine Adverse Event Reporting System

Botsis T, et al

Corresponding author: Taxiarchis Botsis, Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Research (CBER), US FDA, Woodmont Office Complex 1, Rm 306N, 1401 Rockville Pike, Rockville, MD 20852, USA

E-mail: Taxiarchis.Botsis@fda.hhs.gov

Electronic Supplementary Material

Appendix 1: Brighton Collaboration case definition of anaphylaxis

1. Major and minor criteria used in the case definition of anaphylaxis

Organ Systems	Major Criteria
Dermatologic or mucosal	generalized urticaria (hives) or generalized erythema
	angioedema (Not hereditary), localized or generalized
	generalized pruritus with skin rash
Cardiovascular	measured hypotension
	clinical diagnosis of uncompensated shock, indicated by the combination of at least 3 of the following: <ul style="list-style-type: none"> • tachycardia • capillary refill time >3 sec • reduced central pulse volume • decreased level or loss of consciousness
Respiratory	bilateral wheeze (bronchospasm)
	stridor
	upper airway swelling (lip, tongue, throat, uvula, or larynx)
	respiratory distress—2 or more of the following: <ul style="list-style-type: none"> • tachypnoea • increased use of accessory respiratory muscles (sternocleidomastoid, intercostals, etc.) • recession • cyanosis • grunting

Organ Systems	Minor Criteria
Dermatologic or mucosal	generalized pruritus without skin rash
	generalized prickle sensation
	localized injection site urticarial
	red and itchy eyes
Cardiovascular	reduced peripheral circulation as indicated by the combination of at least 2 of: <ul style="list-style-type: none"> • tachycardia and • a capillary refill time of >3 sec without hypotension • a decreased level of consciousness
Respiratory	persistent dry cough
	hoarse voice
	difficulty breathing without wheeze or stridor
	sensation of throat closure
	sneezing, rhinorrhea
Gastrointestinal	diarrhea
	abdominal pain
	nausea
	vomiting
Laboratory	mast cell tryptase elevation > upper normal limit

2. Case definition of anaphylaxis

For all levels of diagnostic certainty
Anaphylaxis is a clinical syndrome characterized by: <ul style="list-style-type: none">• sudden onset AND• rapid progression of signs and symptoms AND• involving multiple (≥ 2) organ systems, as follows:
Level 1 of diagnostic certainty <ul style="list-style-type: none">• ≥ 1 major dermatological AND• ≥ 1 major cardiovascular AND/OR ≥ 1 major respiratory criterion
Level 2 of diagnostic certainty <ul style="list-style-type: none">• ≥ 1 major cardiovascular AND ≥ 1 major respiratory criterion OR• ≥ 1 major cardiovascular OR respiratory criterion AND• ≥ 1 minor criterion involving ≥ 1 different system (<i>other than</i> cardiovascular or respiratory systems) OR• (≥ 1 major dermatologic) AND (≥ 1 minor cardiovascular AND/OR minor respiratory criterion)
Level 3 of diagnostic certainty <ul style="list-style-type: none">• ≥ 1 minor cardiovascular OR respiratory criterion AND• ≥ 1 minor criterion from each of ≥ 2 different systems/categories

MOs try to identify the Brighton Collaboration criteria in each report (i.e. major and minor criteria; section 1 of Appendix 1), fit them into a pattern (i.e. involvement of multiple organ systems; section 2 of Appendix 1) and classify the report as potentially positive or negative in the first step of the review process.

Appendix 2: Brighton Collaboration case definition of GBS

<p>Level 1 of diagnostic certainty</p> <ul style="list-style-type: none">• Bilateral AND flaccid weakness of the limbs AND• Decreased or absent deep tendon reflexes in weak limbs AND• Monophasic illness pattern AND interval between onset and nadir of weakness between 12 h and 28 days AND subsequent clinical plateau AND• Electrophysiologic findings consistent with GBS AND• Cytoalbuminologic dissociation (i.e., elevation of cerebrospinal fluid (CSF) protein level above laboratory normal value AND CSF total white cell count <50 cells/μl) AND• Absence of an identified alternative diagnosis for weakness
<p>Level 2 of diagnostic certainty</p> <ul style="list-style-type: none">• Bilateral AND flaccid weakness of the limbs AND• Decreased or absent deep tendon reflexes in weak limbs AND• Monophasic illness pattern AND interval between onset and nadir of weakness between 12 h and 28 days AND subsequent clinical plateau AND• CSF total white cell count <50 cells/μl (with or without CSF protein elevation above laboratory normal value OR• IF CSF not collected or results not available, electrophysiologic studies consistent with GBS AND• Absence of identified alternative diagnosis for weakness
<p>Level 3 of diagnostic certainty</p> <ul style="list-style-type: none">• Bilateral AND flaccid weakness of the limbs AND• Decreased or absent deep tendon reflexes in weak limbs AND• Monophasic illness pattern AND interval between onset and nadir of weakness between 12 h and 28 days AND subsequent clinical plateau AND• Absence of identified alternative diagnosis for weakness

In summary the following clinical criteria must be present: bilateral and flaccid weakness of the limbs, decreased or absent deep tendon reflexes in weak limbs, a monophasic illness pattern, weakness nadir reached between 12 hours and 28 days, subsequent clinical plateau, and absence of an alternative diagnosis for weakness. Brighton Level 1 (the most stringent level) requires the above clinical criteria AND electrophysiologic findings consistent with GBS, AND the presence of cytoalbuminologic dissociation (elevation of cerebrospinal fluid protein level above laboratory normal value, with CSF total white cell count <50 cells/mm). Brighton Level 2 requires the clinical criteria and EITHER electrophysiologic findings OR the absence of a cerebrospinal fluid pleocytosis (total white cell count <50 cells/mm). Brighton Level 3 requires only the clinical criteria.

Appendix 3: Mapping the BC definition criteria to MedDRA PTs

A. Anaphylaxis; this list was the basis for the first version of the anaphylaxis query.

BC case definition criteria	MedDRA Preferred Terms
accessory respiratory muscles, increased use of	accessory muscle
anaphylaxis	acute anaphyl, acute anaphylactic reaction, acute anaphylaxis, anaphylactic reaction, anaphylactic reaction drug, anaphylactic reaction vaccine, anaphylactic shock, anaphylactic type reaction, anaphylactic vascular purpura, anaphylactoid purpura, anaphylactoid reaction, anaphylactoid shock, anaphylactoid vascular purpura, anaphylaxis, delayed anaphylactoid reaction, purpura anaphylactoid, reaction anaphylactic anaphylactoid, shock anaphylactic anaphylactoid, systemic anaphylactic reaction, systemic anaphylaxis
angioedema, localized or generalized	acute angio edema, acute angio oedema, allergic angioedema, angioedema, angioedema aggravated, angioneurotic edema, angioneurotic edema aggravated, angioneurotic oedema, angioneurotic oedema aggravated, edema angioneurotic, edema vascular, oedema angioneurotic, oedema vascular
consciousness, decreased level of	alertness decreased, arousal difficult, blacked out, blackout, blackout spell, central nervous system depression nos, cns depression, cns depression nos, consciousness abnormal, consciousness awaking loss, consciousness clouding, consciousness decreased, consciousness disturbance transient, consciousness disturbed, consciousness fluctuating, consciousness loss, consciousness loss of, depression central nervous system, falling out, loc, lost consciousness, obnubilation, obtundation, obtundation mental, passed out, precoma, sensorium decreased, twilight state, unconscious, unconscious partial, unconsciousness
cough, persistent, dry	acute cough, allergic cough, chronic cough, cough, cough aggravated, cough decreased, cough decreased, cough increased, cough nonproductive, cough resembling asthma, cough variant asthma, coughing, coughing blood, dry cough, hiccough, irritant cough, irritative cough, nocturnal cough, painful cough, paroxysmal cough, persistent cough, persistent dry cough, productive cough, whooping cough
cyanosis	acrocyanosis, circumoral cyanosis, cyanosis, cyanosis aggravated, cyanosis central, cyanosis central aggravated, cyanosis neonatal, cyanosis nos, cyanosis peripheral
diarrhoea	acute bacterial diarrhea, acute bacterial diarrhoea, acute diarrhea, acute diarrhoea, bacterial diarrhea, bacterial diarrhoea, bloody diarrhea, bloody diarrhoea, chronic diarrhea, chronic diarrhoea, diarrhea, diarrhea aggravated, diarrhea bloody, diarrhea hemorrhagic, diarrhea infectious, diarrhea infectious neonatal, diarrhea neonatal, diarrhea nos, diarrhea recurrent, diarrhoea, diarrhoea aggravated, diarrhoea bloody, diarrhoea haemorrhagic, diarrhoea infectious, diarrhoea infectious neonatal, diarrhoea neonatal, diarrhoea nos, diarrhoea recurrent, explosive diarrhea, explosive diarrhoea, frank diarrhea, frank diarrhoea, haemorrhagic diarrhoea, hemorrhagic diarrhea, iatrogenic diarrhea, iatrogenic diarrhoea, idiopathic diarrhea, idiopathic diarrhoea, infectious diarrhea, infectious diarrhoea, loose bowel, loose bowels, loose motions, loose stools, malodorous diarrhea, malodourous

	diarrhoea, mucous diarrhea, mucous diarrhoea, mushy diarrhea, mushy diarrhoea, mushy stool, nocturnal diarrhea, nocturnal diarrhoea, paradoxical diarrhea, paradoxical diarrhoea, secretory diarrhea, secretory diarrhoea, soft stools, stools loose, stools watery, urgent diarrhea, urgent diarrhoea, viral diarrhea, viral diarrhoea, watery diarrhea, watery diarrhoea
difficulty breathing	breath shortness, breath sounds abnormal, breath sounds absent, breath sounds decreased, breath stacking, breathing abnormally deep, breathing abnormally shallow, breathing arrested, breathing difficult, breathing rate increased, breathing rate slowed, breathing shallow, breathing slowed, breathlessness, difficulty breathing, dyspnoea, hyperventilation, hypopnoea, jaw breathing, labored breathing, laboured breathing, mouth breathing, respiratory arrest, respiratory rate decreased, respiratory rate increased, retractive breathing, shallow breathing, vesicular breathing abnormal
erythema, generalized	application site erythema, application site redness, bullous erythema multiforme, em, epiglottic erythema, erythema, erythema ab igne, erythema annulare, erythema annulare centrifugum, erythema annulare rheumaticum, erythema auricular, erythema bullosum, erythema circumocular-blepharal, erythema diffuse, erythema diutinum, erythema edematous, erythema eyelid, erythema facial, erythema induratum, erythema marginatum, erythema multiforme, erythema multiforme exudativum, erythema multiforme major, erythema multiforme minor, erythema multiforme severe, erythema multiforme type skin reaction, erythema multiforme type vasculitis, erythema nos, erythema oedematous, erythema periorbital, erythema polymorphe, erythematous dermatosis, erythematous conditions, erythematous eruption, eye oedema, eyelid margin hyperaemia, eyelid margin hyperemia, face oedema, face red, flushing, generalised erythema, generalized erythema, genital erythema, gingival erythema, hyperaemia eyelid, hyperemia eyelid, induration erythema, infiltrative erythema, laryngeal erythema, localised erythema, localized erythema, lower legs erythematous, morbilliform erythema, palmar erythema, plantar erythema, red face, red neck, red throat, redness, redness facial, redness generalised, redness generalized, redness gum, redness in breast, redness mouth, skin erythema, skin erythema desquamative, skin hyperaemia, skin hyperemia, skin red, unspecified erythematous condition
eyes, red and itchy	eyes, red and itchy, eye itching, eye pruritus, eye red, eye redness, itching eyes, itchy eyes, ocular hyperaemia, ocular itching
grunting	grunting
hypotension	acute hypotension, arterial hypotension, blood pressure decreased, blood pressure low, chronic hypotension, diastolic hypotension, hypotension, hypotension aggravated, hypotension asymptomatic, hypotension nos, hypotension orthostatic, hypotension orthostatic asymptomatic, hypotension orthostatic symptomatic, hypotension paroxysm, hypotension postural, hypotension postural aggravated, hypotension symptomatic, hypotensive, hypotensive episode, iatrogenic hypotension, low blood pressure, low bp, neonatal hypotension, orthostatic hypotension, postural hypotension, preshock, transient systolic hypotension
nausea	nausea, nausea aggravated, nausea alone, nauseated, nauseous
pain, abdominal	abd pain, abdo pain, abdominal colic, abdominal cramp,

	abdominal cramps, abdominal crampy pains, abdominal pain, abdominal pain aggravated, abdominal pain generalised, abdominal pain generalized, abdominal pain localised, abdominal pain localized, abdominal pain lower, abdominal pain nos, abdominal pain peptic ulcer type, abdominal pain upper, ache stomach, acute gastric pain, belly ache, bowel cramps, bowel spasm, central abdominal pain, chronic abdominal pain, chronic epigastric pain, colic, colic abdominal, colicky, cramp abdominal, cramp in lower abdomen, epigastralgia, epigastric ache, epigastric pain, epigastric pain epigastralgia, gastralgia, gastric pain, gastric spasm, gastrointestinal cramps, gastrointestinal pain, gastrointestinal pain nos, gastrointestinal stinging, gi pain, gut pain, hypogastric pain, intestinal cramps, intestinal spasm, intestine burning sensation of, left lower quadrant pain, left upper quadrant pain, lower abdominal pain, pain abdo, pain abdominal, pain epigastric, pain gastric, pain hunger, pain right upper quadrant, pain right upper quadrant, pain stomach, pancreatic pain, peritoneal pain, periumbilical pain, right lower quadrant pain, right upper quadrant pain, ruq pain, spasm intestinal, splenic flexure syndrome, stomach ache, stomach cramps, stomach pain, stomachache, ulcer type pain, upper abdominal pain
prickle sensation	burning sensation, choking, choking sensation
pruritus, generalized	brachioradial pruritus, ear pruritus, eyeball itching, eyelids itchy sensation of, eyelids pruritus, general pruritus, generalised itching, generalized itching, generalized pruritus, genital itching, genital itching female, genital itching male, genital itching nos, genital pruritus female, genital pruritus male, genital pruritus nos, gingival pruritus, hairy skin itching, intraoral pruritus, itch, itch burning, itching, itching gum, itching mouth, itchy, itchy legs, itchy scalp, itchy skin, lip pruritus, oral pruritus, other specified pruritic conditions, pruritis, pruritus, pruritus aggravated, pruritus allergic, pruritus and related conditions, pruritus breast, pruritus cutaneous, pruritus facial, pruritus generalised, pruritus generalized, pruritus nos, unspecified pruritic disorder
pulse volume, reduced central	pulse volume decreased
reduced peripheral circulation	collapse circulatory peripheral, collapse peripheral circulatory, collapse peripheral vascular, failure peripheral circulatory, peripheral circulatory failure, peripheral shutdown, poor peripheral circ, poor peripheral circulation, poor peripheral perfusion
respiratory distress	acute respiratory distress syndrome, adult respiratory distress syndr, adult respiratory distress syndrome, infant respiratory distress syndrome, neonatal respiratory distress, neonatal respiratory distress syndrome, respiratory distress, respiratory distress neonatal, respiratory distress syndrome, respiratory distress syndrome adult, respiratory distress syndrome in newborn, respiratory distress syndrome newborn, syndrome adult respiratory distress, syndrome respiratory distress adult, syndrome respiratory distress newborn
rhinorrhoea	mucous nasal increased, mucus nasal increased, nasal discharge, nasal discharge watery excessive, nasal mucous increased, nasal mucus increased, rhinorrhea, rhinorrhoea, runny nose, sniffles

skin rash	<p>blotchy rash, butterfly rash, dermatitis papular, diffuse macular rash, dry rash face, dry rash neck, eruption, eruption facial, erythematous florid universal rash, erythematous rash, erythematous skin rash, erythromacular rash, erythropapular rash, exanthem, exanthema, exanthema facial, exanthema generalised, exanthema generalized, exfoliative rash, facial rash, generalised eruption, generalised macular rash, generalised raised red spots limbs, generalised raised red spots trunk, generalised rash, generalized eruption, generalized macular rash, generalized raised red spots limbs, generalized raised red spots trunk, generalized rash, hand rash, heat rash, hypersensitive syndrome, hypersensitivity syndrome, itchy rash, localised rash, localized rash, macular rash, maculo-papular, maculopapular rash, malar rash, measly rash, morbilliform eruption, morbilliform papular rash, morbilliform rash, mucocutaneous rash, neck rash, nettle rash, palmoplantar rash, papular rash, papular skin eruption, papulopustular rash, papulosquamous rash, papulovesicular rash, patchy rash, periocular rash, prickly heat, pruritic rash, pustular rash, pustular skin eruption, pustule, rash, rash aggravated, rash desquamating, rash erythematous, rash erythematous aggravated, rash face, rash follicular, rash generalised, rash herpeticiform, rash macular, rash maculovesicular, rash measly, rash morbilliform, rash mouth, rash neonatal, rash nos, rash over arms, rash papular, rash papulosquamous, rash papulovesicular, rash peripheral, rash pruritic, rash pustular, rash recurrent, rash rubelliform, rash scaly, rash scarlatiniform, rash trunk, rash varicelliform, rash vesicular, rash vesiculopustular, rash zosteriform, red blotches, red confluent rash, red confluent rash both legs, red confluent rash cheeks, septic rash, septic spots, skin blotches, skin eruption, skin rash, sle rash, sweat rash, systemic erythematosus rash, systemic lupus erythematosus rash, vasculitic rash, vesicular eruption, vesicular rash, vesiculopustular rash, viral exanthem, viral rash, viral rash nos, widespread rash</p>
sneezing	Sneezing, paroxysmal sneeze, sneezing excessive
stridor	laryngospasm, stridor, stridor inspiratory
swelling, upper airways	<p>laryngeal oedema, lip oedema, lip swelling, nasal congestion, nasal mucosal swelling, oedema mouth, oral mucosa swollen, oropharyngeal spasm, oropharyngeal swelling, palatal oedema, pharyngeal oedema, swelling lips, swollen lips, swollen mouth, swollen tongue, throat swelling, throat swelling nos, tongue red swollen pain</p>
tachycardia	<p>acute supraventricular tachycardia, atrial tachycardia, atrioventricular nodal reentrant tachycardia, av nodal reentrant tachycardia, av nodal tachycardia, av reentrant tachycardia, fetal supraventricular tachycardia, fetal tachycardia, foetal supraventricular tachycardia, foetal tachycardia, incessant ventricular tachycardia, monomorphic ventricular tachycardia, multifocal atrial tachycardia, multifocal ventricular tachycardia, multiform ventricular tachycardia, neonatal tachycardia, paroxysmal atrial tachycardia, paroxysmal supraventricular tachycardia, paroxysmal ventricular tachycardia, paroxysmal vt, pat, polymorphic ventricular tachycardia, postural orthostatic tachycardia syndrome, rebound tachycardia, recurrent sustained ventricular tachycardia, recurrent ventricular tachycardia, reflex tachycardia, sinus tachycardia, supraventricular tachycardia, sustained ventricular tachycardia, svt, tachycardia, tachycardia</p>

	aggravated, tachycardia atrial, tachycardia nervous, tachycardia nos, tachycardia paroxysmal, tachycardia paroxysmal atrial, tachycardia paroxysmal nos, tachycardia paroxysmal ventricular, tachycardia sinus, tachycardia supraventricular, tachycardia supraventricular aggr, tachycardia ventricular, ventricular tachycardia, vt, wide complex supraventricular tachycardia, wide complex tachycardia, wide complex ventricular tachycardia
tachypnoea	neonatal tachypnoea, neonatal tachypnea, tachypnoea, tachypnea, polypnea, polypnoea, respiration stimulated
throat closure, sensation of	burning in throat, constriction throat, dry throat, itchy throat, numbness throat, sore throat, sore throat nos, throat burning sensation of, throat constriction, throat dry, throat edema, throat infection, throat irritation, throat lesion, throat oedema, throat pain, throat sore, throat tightness
uncompensated shock	allergic shock, cardiogenic shock, cardioversion, circulatory collapse, clinically shocked, dc shock, distributive shock, shock, shock symptom, shock symptoms, shock vascular, vasovagal shock
urticaria (hives), generalized	acute allergic urticaria, acute urticaria, allergic urticaria, erythema urticarial, generalized urticaria, generalized urticarial rash, giant hives, giant urticaria, haemorrhagic urticaria, hemorrhagic urticaria, hives, hives giant, idiopathic urticaria, other specified urticaria, papular urticaria, papular urticarial eruption, unspecified urticaria, urticaria, urticaria acute, urticaria aggravated, urticaria generalised, urticaria giant, urticaria haemorrhagica, urticaria hemorrhagica, urticaria idiopathic, urticaria injection site, urticaria localised, urticaria localized, urticaria nos, urticaria papular, urticaria pigmentosa, urticaria vesiculosa, urticarial, urticarial rash, urticarial symptom, urticarial vasculitis, urticated rash, urtication, welts, wheals
voice, hoarse	aphonia, distorted voice, dysphonia, hoarse voice, other voice disturbance, voice alteration, voice disturbance, voice lowered, voicelessness
vomiting, unspecified	bilious vomiting, blennemesis, coffee ground vomiting, emesis, faecal vomiting, fecal vomiting, habit vomiting, haematemesis, hyperemesis, other vomiting complicating pregnancy, persistent vomiting, vomited, vomiting, vomiting aggravated, vomiting alone, vomiting blood, vomiting in pregnancy, vomiting neonatal, vomiting nos, vomiting projectile, vomiting reflex
wheezing(bronchospasm), bilateral	allergic bronchospasm, asthmatic wheezing, bronchial spasm, bronchial stricture, bronchoconstriction, bronchospasm, bronchospasm aggravated, bronchospasm nos, bronchospasm paradoxical, increased wheeziness, paradoxical bronchospasm, prophylaxis against bronchospasm, spasm bronchial, wheeze, wheeze worsened, wheezes, wheezing, wheezing aggravated, wheezing expiratory, wheezing inspiratory, wheezy, wheezy bronchitis

B. GBS; this list was the basis for the first version of the GBS query.

BC case definition criteria	MedDRA Preferred Terms
CSF protein level above laboratory normal value	CSF protein increased, CSF immunoglobulin increased, CSF oligoclonal band present, CSF protein abnormal, CSF test abnormal, Albumin CSF increased
CSF total white cell < 50 cell/mm ³	Lumbar puncture abnormal

Decreased or absent, bilateral deep tendon reflexes	Areflexia, Reflexes abnormal, Hyporeflexia, Autonomic dysreflexia
Electrophysiologic findings consistent with GBS	Electromyogram abnormal, Nerve conduction studies abnormal, Nerve stimulation test abnormal
Identified alternative diagnosis of weakness other than GBS	Cardiac autonomic neuropathy, Central nervous system inflammation, Central nervous system lesion, Deafness permanent, Deafness transitory, Lower motor neurone lesion, Myelopathy, Nerve injury, Nervous system disorder, Neuritis, Neuritis cranial, Neuropathy peripheral, Peripheral motor neuropathy, Peripheral nerve injury, Peripheral nerve lesion, Peripheral nerve palsy, Peripheral nerve paresis, Peripheral paralysis, Polyneuropathy idiopathic progressive, Vocal cord paralysis, Vocal cord paresis, Accessory nerve disorder, Spastic paralysis, Spastic paraplegia, Sympathetic nerve injury, Upper motor neurone lesion, Vagus nerve disorder
Weakness of limbs, bilateral	Acoustic neuritis, Autoimmune neuropathy, Autonomic failure syndrome, Autonomic neuropathy, Axonal neuropathy, Cranial nerve injury, Cranial nerve palsies multiple, Cranial nerve paralysis, Deafness neurosensory, Demyelinating polyneuropathy, Demyelination, Gait apraxia, Gait deviation, Gait disturbance, Glossopharyngeal nerve disorder, Glossopharyngeal nerve paralysis, Glossopharyngeal neuralgia, Hypoglossal nerve disorder, Hypoglossal nerve paralysis, Hypoglossal nerve paresis, IIIrd nerve disorder, IIIrd nerve injury, IIIrd nerve paralysis, IIIrd nerve paresis, IVth nerve disorder, IVth nerve injury, IVth nerve paralysis, IVth nerve paresis, Muscular weakness, Oculofacial paralysis, Olfactory nerve disorder, Ophthalmoplegia, Optic nerve disorder, Optic nerve injury, Optic neuropathy, Polyneuropathy, Trigeminal palsy, VIIIth nerve injury, VIIIth nerve lesion, VIIth nerve injury, VIIth nerve paralysis, Visual field defect, VIth nerve disorder, VIth nerve injury, VIth nerve paralysis, VIth nerve paresis, Vth nerve injury, XIIth nerve injury, XIth nerve injury, XIth nerve paralysis, Respiratory muscle weakness, Eye movement disorder, Eyelid ptosis, Facial nerve disorder, Facial paresis
Weakness of limbs, flacid	Paralysis, Paralysis flaccid, Paraparesis, Paraplegia, Paresis, Paresis anal sphincter, Paresis cranial nerve, Phrenic nerve paralysis, Quadriparesis, Quadriplegia, Respiratory paralysis, Vagus nerve paralysis
GBS related terms	Acute polyneuropathy, Chronic inflammatory demyelinating polyradiculoneuropathy, Guillain-Barre syndrome, Miller Fisher syndrome

Appendix 4: Cosine similarity calculations (Vector Space Model)

The first version of the anaphylaxis query vector $\vec{V}(q_{BC_anaph})$ consisted of all the medical terms in the major and minor criteria of the BC case definition as well as the ‘anaphylaxis term’, while the report vector $\vec{V}(r_{BC_anaph})$ consisted of the corresponding extracted medical terms. Thus, we created a query vector of 32 components:

$\vec{V}(q_{BC_anaph}) = [$ ‘accessory respiratory muscles, increased use of’ ‘anaphylaxis’ ‘angioedema, localized or generalized’ ‘consciousness, decreased level of’ ‘cough, persistent, dry’ ‘cyanosis’ ‘diarrhoea’ ‘difficulty breathing’ ‘erythema, generalized’ ‘eyes, red and itchy’ ‘grunting’ ‘hypotension’ ‘nausea’ ‘pain, abdominal’ ‘prickle sensation’ ‘pruritus, generalized’ ‘pulse volume, reduced central’ ‘reduced peripheral circulation’ ‘respiratory distress’ ‘rhinorrhoea’ ‘skin rash’ ‘sneezing’ ‘stridor’ ‘swelling, upper airways’ ‘tachycardia’ ‘tachypnoea’ ‘throat closure, sensation of’ ‘uncompensated shock’ ‘urticaria (hives), generalized’ ‘voice, hoarse’ ‘vomiting, unspecified’ ‘wheezing(bronchospasm), bilateral’]

An example of a report vector $\vec{V}(r_{BC_anaph})$ with 6 anaphylaxis-related components:

$\vec{V}(r_{BC_anaph}) = [$ ‘erythema, generalized’ ‘eyes, red and itchy’ ‘pruritus, generalized’ ‘sneezing’ ‘throat closure, sensation of’ ‘urticaria (hives), generalized’]

The second version of the anaphylaxis query included the exact 83 MedDRA PTs of the anaphylaxis SMQ:

$\vec{V}(q_{smq_anaph}) = [$ ‘anaphylactic reaction’ ‘anaphylactic shock’ ‘anaphylactic transfusion reaction’ ‘anaphylactoid reaction’ ‘anaphylactoid shock’ ‘circulatory collapse’ ‘first use syndrome’ ‘kounis syndrome’ ‘shock’ ‘type i hypersensitivity’ ‘acute respiratory failure’ ‘asthma’ ‘bronchial oedema’ ‘bronchospasm’ ‘cardio’ ‘respiratory distress’ ‘chest discomfort’ ‘choking’ ‘choking sensation’ ‘circumoral oedema’ ‘cough’ ‘dyspnoea’ ‘hyperventilation’ ‘laryngeal dyspnoea’ ‘laryngeal oedema’ ‘laryngospasm’ ‘laryngotracheal oedema’ ‘nasal obstruction’ ‘oedema mouth’ ‘oropharyngeal spasm’ ‘oropharyngeal swelling’ ‘respiratory arrest’ ‘respiratory distress’ ‘respiratory failure’ ‘reversible airways obstruction’ ‘sensation of foreign body’ ‘sneezing’ ‘stridor’ ‘swollen tongue’ ‘tachypnoea’ ‘throat tightness’ ‘tongue oedema’ ‘tracheal obstruction’ ‘tracheal oedema’ ‘upper airway obstruction’ ‘wheezing’ ‘allergic oedema’ ‘angioedema’ ‘cyanosis’ ‘erythema’ ‘eye oedema’ ‘eye pruritus’ ‘eye swelling’ ‘eyelid oedema’ ‘face oedema’ ‘fixed eruption’ ‘flushing’ ‘generalised erythema’ ‘injection site urticaria’ ‘lip oedema’ ‘lip swelling’ ‘ocular hyperaemia’ ‘oedema’ ‘periorbital oedema’ ‘pruritus’ ‘pruritus allergic’ ‘pruritus generalised’ ‘rash’ ‘rash erythematous’ ‘rash generalised’ ‘rash pruritic’ ‘skin swelling’ ‘swelling’ ‘swelling face’ ‘urticaria’ ‘urticaria papular’ ‘blood pressure decreased’ ‘blood pressure diastolic decreased’ ‘blood pressure systolic decreased’ ‘cardiac arrest’ ‘cardio’ ‘respiratory arrest’ ‘cardiovascular insufficiency’ ‘diastolic hypotension’ ‘hypotension’]

The vector for the above report $\vec{V}(r_{SMQ_anaph})$ included only the anaphylaxis SMQ-related terms, a total of 5 components, instead of 6 in the BC-based vector (‘throat closure, sensation of’ is not included in the anaphylaxis SMQ):

$\vec{V}(r_{smq_anaph}) = [\text{'Erythema' 'Eye pruritus' 'Ocular hyperaemia' 'Sneezing' 'Urticaria'}]$

Similarly, the first and second version of the GBS query vector consisted of 8 and 118 components, respectively:

$\vec{V}(q_{BC_gbs}) = [\text{'CSF protein level above laboratory normal value' 'CSF total white cell < 50 cell/mm}^3\text{'}$
 $\text{'decreased or absent, bilateral deep tendon reflexes' 'Electrophysiologic findings consistent with GBS' 'identified alternative diagnosis of weakness other than GBS' 'Weakness of limbs, bilateral' 'Weakness of limbs, flacid'}$
 $\text{'GBS related terms'}]$

$\vec{V}(q_{smq_gbs}) = [\text{'chronic inflammatory demyelinating polyradiculoneuropathy' 'demyelinating polyneuropathy'}$
 $\text{'guillain-barre syndrome' 'miller fisher syndrome' 'abasia' 'acute respiratory distress syndrome' 'akinaesthesia'}$
 $\text{'albumin csf increased' 'aphasia' 'areflexia' 'asthenia' 'ataxia' 'autoimmune neuropathy' 'autonomic nervous system imbalance'}$
 $\text{'autonomic neuropathy' 'axonal neuropathy' 'balance disorder' 'bell's phenomenon' 'biopsy peripheral nerve abnormal'}$
 $\text{'brain stem auditory evoked response abnormal' 'bulbar palsy' 'central nervous system function test abnormal'}$
 $\text{'central nervous system lesion' 'central nervous system mass' 'coordination abnormal' 'cranial nerve disorder'}$
 $\text{'cranial nerve palsies multiple' 'cranial nerve paralysis' 'csf immunoglobulin increased' 'csf oligoclonal band present'}$
 $\text{'csf protein abnormal' 'csf protein increased' 'csf test abnormal' 'decreased vibratory sense' 'demyelination'}$
 $\text{'diaphragmatic paralysis' 'diplegia' 'dysaesthesia' 'dysarthria' 'dysphagia' 'dysstasia'}$
 $\text{'electromyogram abnormal' 'extensor plantar response' 'facial paresis' 'gait disturbance' 'genital hypoaesthesia'}$
 $\text{'hypoaesthesia' 'hypoaesthesia eye' 'hypoaesthesia oral' 'hypoaesthesia teeth' 'hypokinesia' 'hyporeflexia'}$
 $\text{'hyporesponsive to stimuli' 'hypotonia' 'intranasal hypoaesthesia' 'intranasal paraesthesia' 'loss of proprioception'}$
 $\text{'lumbar puncture abnormal' 'motor dysfunction' 'multifocal motor neuropathy' 'muscle atrophy' 'muscular weakness'}$
 $\text{'nerve conduction studies abnormal' 'nerve degeneration' 'nerve stimulation test abnormal' 'neuromyopathy'}$
 $\text{'neuronal neuropathy' 'neuropathy peripheral' 'ophthalmoplegia' 'papilloedema' 'paraesthesia' 'paraesthesia mucosal'}$
 $\text{'paraesthesia of genital female' 'paraesthesia of genital male' 'paraesthesia oral' 'paralysis' 'paralysis flaccid'}$
 $\text{'paraparesis' 'paraplegia' 'paresis' 'paresis cranial nerve' 'peripheral motor neuropathy' 'peripheral nerve lesion'}$
 $\text{'peripheral nerve palsy' 'peripheral nervous system function test abnormal' 'peripheral paralysis' 'peripheral sensorimotor neuropathy'}$
 $\text{'peripheral sensory neuropathy' 'peroneal nerve palsy' 'pharyngeal hypoaesthesia' 'polyneuropathy' 'polyneuropathy idiopathic progressive'}$
 $\text{'progressive bulbar palsy' 'pseudobulbar palsy' 'quadriparesis' 'quadriplegia' 'radiculopathy'}$
 $\text{'reflexes abnormal' 'respiratory arrest' 'respiratory failure' 'respiratory muscle weakness' 'respiratory paralysis'}$
 $\text{'sensorimotor disorder' 'sensory disturbance' 'sensory integrative dysfunction' 'sensory level abnormal'}$
 $\text{'sensory loss' 'speech disorder' 'spinocerebellar disorder' 'tandem gait test abnormal' 'trigeminal nerve paresis'}$
 $\text{'trigeminal palsy' 'ulnar neuritis' 'unresponsive to stimuli' 'vibration test abnormal' 'viith nerve paralysis'}$
 $\text{'vith nerve paralysis' 'vith nerve paresis'}]$

The corresponding example of the BC-based report vector $\vec{V}(r_{GBS})$ included 4 components:

$\vec{V}(r_{BC_gbs}) = [\text{'Decreased or absent, bilateral deep tendon reflexes' 'Electrophysiologic findings consistent with GBS'}$
 $\text{'GBS related terms' 'Weakness of limbs, bilateral'}]$

The SMQ-based report vector included 11 components:

$\vec{V}(r_{smq_gbs}) = [\text{'guillain-barre syndrome' 'areflexia' 'gait disturbance' 'muscular weakness' 'paraesthesia' 'hyporeflexia' 'sensory loss' 'hypoesthesia' 'abasia' 'demyelinating polyneuropathy' 'electromyogram abnormal'}]$

According to Appendix 3, the components in the two report vectors could be mapped as follows: (i) 'reflexia' and 'hyporeflexia' to 'Decreased or absent, bilateral deep tendon reflexes'; (ii) 'electromyogram abnormal' to 'Electrophysiologic findings consistent with GBS'; (iii) 'guillain-barre syndrome' to 'GBS related terms'; and (iv) 'gait disturbance', 'muscular weakness' and 'demyelinating polyneuropathy' to 'Weakness of limbs, bilateral'. 'Paraesthesia', 'sensory loss', 'hypoesthesia' and 'abasia' were not included in the BC-related terms by the medical expert and could not be mapped to the BC broad categories.

IR suggests the formulation of numerical vectors through the assignment of certain weights to the vector components. To formulate the numerical query vector we calculated the weight of each component that was equal to the inverse report frequency (irf_t) of the corresponding term in the training set or equal to zero when no reports included this term. The irf_t (or the weight of each component in the query – $w_{t,q}$) was equal to:

$$w_{t,q} = \log \frac{N}{rf_t} \quad (1)$$

where N was the number of reports in the training set and rf_t the number of reports that contained the term 't'. Subsequently the $w_{t,q}$ was length-normalized divided by the Euclidean length of the query vector, e.g. the normalized weight for 'areflexia' included in the second version of the GBS query the was:

$$normalized_w_{t_{areflexia},q_{smq_gbs}} = \frac{w_{t_{areflexia},q_{smq_gbs}}}{\sqrt{(w_{t_1,q_{smq_gbs}})^2 + (w_{t_2,q_{smq_gbs}})^2 + \dots + (w_{t_{118},q_{smq_gbs}})^2}} \quad (2)$$

where t_1 , t_2 and t_{118} were equal to 'chronic inflammatory demyelinating polyradiculoneuropathy', 'demyelinating polyneuropathy' and 'vith nerve paresis', respectively. 'Areflexia' was the 10th component of the query vector, so it could alternatively appear as t_{10} in the above equation.

For the report vectors, we used the rf_t weighting with no use of irf_t however with Euclidean normalization. Thus the normalized weight of 'areflexia' in the above SMQ-based report was:

$$normalized_w_{t_{areflexia},r_{smq_gbs}} = \frac{w_{t_{areflexia},r_{smq_gbs}}}{\sqrt{(w_{t_1,r_{smq_gbs}})^2 + \dots + (w_{t_{118},r_{smq_gbs}})^2}} \quad (3)$$

The cosine similarity (or score) for the above SMQ-based GBS report was equal to:

$$sim(r_{smq_gbs}, q_{smq_gbs}) = \sum \left(normalized_{w_{t_1, q_{smq_gbs}}} \times normalized_{w_{t_1, r_{smq_gbs}}} \right) + \dots \quad (4)$$

$$+ \left(normalized_{w_{t_{118}, q_{smq_gbs}}} \times normalized_{w_{t_{118}, r_{smq_gbs}}} \right)$$

The subscripts denote the position of the components in the query and report vectors.

It should be also noted that the above equation is equivalent to the general equation used for the calculation of the cosine similarity of the report vs. the GBS query numerical vector:

$$sim(r_{smq_gbs}, q_{smq_gbs}) = \frac{\vec{V}(r_{smq_gbs}) \cdot \vec{V}(q_{smq_gbs})}{|\vec{V}(r_{smq_gbs})| \cdot |\vec{V}(q_{smq_gbs})|} \quad (5)$$

where the numerator represents the dot product of the report and query vectors, while the denominator their Euclidean lengths.

It should be clarified that the component ‘identified alternative diagnosis of weakness other than GBS’ that was included in the first version of the GBS query vector and, potentially, in any report vector of the GBS sample corresponds to the following criterion in the BC case definition for GBS: ‘absence of an identified alternative diagnosis for weakness’. Practically, this component had a negative value in the numerical query vector and, thus, contributed negatively to the total score of the report that included an alternative diagnosis for weakness, such as ‘spastic paralysis’.