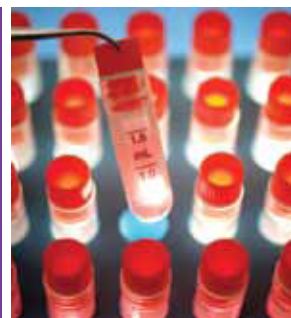


BIOMARKERS MODULE

FREQUENTLY ASKED QUESTIONS



CONTENT

WHAT CAN I FIND IN THE *BIOMARKERS MODULE*?

The *Biomarkers Module* of *Thomson Reuters Integrity*SM covers all key biomarker uses at every stage of drug R&D including disease risk detection, diagnosis, target identification, proof-of-mechanism, proof-of-concept, treatment/safety monitoring, and outcome measurement.

The initial focus of the *Biomarkers Module* was the major therapy areas of: oncology; cardiovascular disease; diabetes; immunology; respiratory disorders and neurological disorders. Therapy areas are continually being expanded and now proactively cover, pain, infectious disease, musculoskeletal and connective tissue disorders and endocrine disease.

For each therapy area, the *Biomarkers Module* includes not only well-established biomarkers but also high potential biomarkers as identified by industry and academia experts.

The database covers a variety of types of biomarkers including genomic, proteomic, biochemical and cellular (see the 'DEFINITIONS & ABBREVIATIONS' section of this FAQ document). In January 2010, we began covering imaging markers in cardiovascular, respiratory and neurological disorders, including structural and physiological biomarkers. From February 2011, panel or combination biomarkers are included with the initial focus on genomic cancer panels.

HOW DO YOU SELECT BIOMARKERS TO INCLUDE IN THE *BIOMARKERS MODULE*?

We have extensive editorial resources monitoring meetings, literature and patents for biomarkers using a largely manual process. Some of the key players in the process include:

- key opinion leaders who suggest new biomarkers to add to the database based on their up-to-date knowledge of the field or therapy area.
- readers who monitor newly published journals and identify articles relevant to each knowledge area of *Integrity*. Articles about biomarkers are then routed to the *Biomarkers Module* editorial team. A parallel process occurs with the other sources that we cover.
- members of the *Biomarkers Module* Quality Control team whose responsibilities include reading review articles looking for biomarkers.

- a specialized team that search databases for additional information (such as further studies to support or refute a biomarker use) once new biomarkers for the database have been identified. The documents that they retrieve are analyzed by a team of editorial analysts (these documents sometimes contain information on additional biomarkers that are in turn passed on to the QC team for evaluation).

WHAT ARE THE SELECTION CRITERIA TO INCLUDE A BIOMARKER IN THE *BIOMARKERS MODULE*?

There must be reliable scientific evidence in a published source to support a specific use of a biomarker. The minimum content for addition to the database is biomarker name and at least one use. Each use must be associated with a disease, adverse event, underlying disease pathology or a drug target and must have a role (diagnosis, predicting treatment efficacy, etc).

WHO CREATES THE *BIOMARKERS MODULE*?

There is a dedicated team of scientific analysts working exclusively in the biomarkers area. Their educational backgrounds include degrees in medicine, molecular biology and biochemistry. A separate team of PhD qualified scientists is responsible for quality control. We also benefit from expert input from analysts working in the areas of genes, patents, kits, selection of literature, companies and markets. The full team is around 150 people.

WHAT IS THE ROLE OF KEY OPINION LEADERS IN THE SELECTION AND RECOMMENDATION PROCESS?

The key opinion leaders provide guidance on which biomarkers are appropriate for inclusion in the database and details on the uses being studied, as well as suggestions on information sources (for example new congresses — which ones are most worthwhile for content, which ones are more political in nature, etc).

Thomson Reuters has a long tradition of working with key opinion leaders to support product and content development and this approach has been extended to the area of biomarkers.



WHO ARE THE KEY OPINION LEADERS?

The names of some of the current key opinion leaders involved in the *Biomarkers Module* are listed below. Their role is to seed the database, then on a regular basis to identify new biomarkers and/or uses to add to the database. They also provide content and product advice as the database evolves.

Professor Jeffrey S. Ross, M.D.
(Oncology)

Professor Juan Carlos Kaski
(Cardiovascular Science)

Dr Trevor Hansel
(Respiratory Disease)

HOW FAR BACK DOES THE INFORMATION GO IN THE BIOMARKERS MODULE?

When a biomarker is identified in one of the main therapy areas and selected for inclusion in the *Biomarkers Module* the editorial analysts search literature from the last five years and clinical guidelines from the previous 10 years.

Due to the fast rate of biomarker research, if a biomarker is relevant, it will be mentioned in papers published during the last five years. Guidelines can be current but not necessarily recently updated, hence the search is expanded to track older (but still current) guidelines.

WHAT ARE THE INFORMATION SOURCES FOR THE BIOMARKERS MODULE?

Principal information sources are biomedical literature and congresses, clinicaltrials.gov, the FDA website and patent literature (WO, US, EP, JP, CN, KR, and IN).

In more detail, the main sources are:

- **Scientific meetings:** Information is taken from meetings that specialize in biomarkers content (such as World Biomarkers Congress) as well as relevant meetings in all therapeutic areas where data related to biomarkers can also be presented (ASCO, ACCR, ADA, AUA, etc).
- **Literature Information:** References are identified by searching MEDLINER and full-text of peer reviewed journals included in the Web of Science. Guidelines from scientific societies are usually included here.
- **Patents:** From seven patent authorities: Japan Patent Office (JP), World Intellectual Property Organization (WIPO), US Patents & Trademark Office (US), European Patent Office (EP), Chinese patents (CN), Korean patents office (KR) and Indian patents (IN).

WHAT ARE THE UPDATE PROCESS AND GUIDELINES?

The database is updated on a daily basis. Our tracking of information sources not only looks for content on biomarkers not yet entered in the database, but also for new or updated information on biomarkers for which there is an existing record. This could be:

- a new use under study (a use is a combination of indication, population, role and technique)
- information indicating that the use has advanced (for example, gone from experimental use to use in humans, similar to a phase change for a drug)
- a relevant new reference to add to the database, albeit one that doesn't prompt a change in a lifecycle status for the biomarker (or information on a new kit in development, as another example).

WHAT IS THE QUALITY ASSURANCE PROCESS FOR DATA ENTERED INTO THE BIOMARKERS MODULE?

The database is curated by scientists who read the information found in the references, analyze this information and prepare the *Biomarkers Module* record, including the information on the biomarker itself and its uses, as well as on the development of any FDA-approved kits associated with a use (depending on the biomarker's lifecycle stage, of course). They are supported by modern, computerized tools to locate literature and patent references and to process the information.

All records must pass through a rigorous quality control process before being released into the database. The quality control team is independent of the editorial team. Members of the quality control team, with specific expertise, check individual fields of a record. Then the entire record is checked to ensure all the information is internally consistent.

WHAT ARE THE CRITERIA FOR ASSIGNING A REFERENCE TO THE METHODS SECTION OF A BIOMARKERS MODULE RECORD?

From January 2010, a reference (research paper, patent, review article or guideline) is assigned to the Methods section when the objective of the study is to develop or evaluate new methods for measuring the biomarker or to describe how to measure more established biomarkers.

WHAT ARE THE CRITERIA FOR LINKING A REFERENCE TO THE REVIEWS SECTION OF A RECORD IN THE BIOMARKERS MODULE?

Guidelines (from 2000 onwards) which specifically name the biomarker, and reviews (from January 2010) which focus specifically on the biomarker, are linked to the "reviews" section of a biomarker record.

WHAT ARE THE CRITERIA FOR SHOWING A REFERENCE OR PATENT IN THE "SOURCE" COLUMN OF A BIOMARKERS MODULE RECORD?

If, upon analysis, a document (literature reference, clinical practice guideline or patent document) is found to provide highly relevant information to support a specific use of a biomarker, then the document will be added to the "Source" column.

WHY DOES THE NUMBER OF REFERENCES IN "RELATED INFORMATION" DIFFER FROM THE NUMBER OF REFERENCES IN THE "SOURCE" COLUMN?

As mentioned above, a reference in the "Source" column is directly relevant to a specific use of a biomarker. "Related Information" includes all these source documents as well as other, more general references (e.g. reviews) that are relevant to the biomarker but which do not meet the criteria to be associated with a specific use. Some of these articles will be linked to new uses established at a later date.

WHAT ARE THE CRITERIA FOR DESIGNATING A SOURCE REFERENCE OR PATENT AS SUPPORTING OR REFUTING A BIOMARKER USE?

Source references and patents in the *Biomarkers Module* can be classified as supporting, conflicting or mixed.

Supporting (+): A reference/patent is classified as "supporting" when a significant association between biomarker and condition/toxicity/pathology is found ($p \leq 0.05$).

Conflicting (-): A reference/patent is classified as "conflicting" when no significant association is found between the biomarker and condition/toxicity/pathology ($p > 0.05$).

Mixed (+/-): A reference/patent is classified as "mixed" when, in a single source document there is both supporting and conflicting evidence for the biomarker in that use. For example, more than one polymorphism in a gene has been reported in the same study, and only some support the use.

BIOMARKER FIELDS

WHERE DO THE SYNONYMS IN A BIOMARKERS MODULE RECORD COME FROM?

Synonyms are taken from the NCBI, Uniprot, MeSH® terms and Expaty database.

WHAT ARE THE DEFINITIONS FOR TYPES IN THE BIOMARKERS MODULE?

Anthropomorphic biomarkers: are of the body shape/form, for example body mass index.

Cellular biomarkers: are whole cells, for example cancer cells identified by the Pap test.

Biochemical biomarkers: are substrates or products of chemical reactions in living organisms, for example cholesterol and bilirubin.

Genomic biomarkers: are variants in the DNA sequence or in the transcription level, for example HER2.

Physiological biomarkers: are body processes, for example systolic blood pressure.

Proteomic biomarkers: are variants in protein sequence, protein levels in a given tissue, protein interactions, and enzyme activities.

Structural biomarkers: are anatomical structures, e.g. hippocampus, or they are lesions, e.g. atherosclerotic plaque.

HOW IS A BIOLOGICAL PROCESS TERM ASSIGNED TO A BIOMARKERS MODULE RECORD?

When an editorial analyst finds evidence in a source document to support a link between a biomarker and a particular biological process, the process is indexed. The most appropriate (usually the most specific but always the most descriptive) term from the Gene Ontology Consortium list is intellectually assigned based on information in the source.

IS IT POSSIBLE TO SEARCH ON FREE TEXT?

A search in the description field is a free text search. Other fields use controlled vocabulary, rather than free text, to allow more accurate search and retrieval. It is recommended that users select controlled search terms from the browse Index.

WHAT ARE THE CRITERIA TO LINK A MECHANISM MODIFIER TO A BIOMARKERS MODULE RECORD?

A mechanism modifier is linked to a biomarker when the biomarker is a target for a drug.

USE RECORD FIELDS

WHAT IS A USE ID?

A Use ID is a unique serial number assigned to every Use record found in the *Biomarkers Module*.

INDICATION – CONDITION

WHICH THERAPY AREAS ARE COVERED BY THE BIOMARKERS MODULE?

The initial focus of the *Biomarkers Module* was the major therapy areas of: oncology; cardiovascular disease; diabetes; immunology; respiratory disorders and neurological disorders. Therapy areas are continually being expanded and now proactively cover, pain, infectious disease, musculoskeletal and connective tissue disorders and endocrine disease. In each of these core therapeutic areas, information from the last 5 years on all uses of a biomarker is analysed.

WHY DO I SEE INDICATIONS OUTSIDE THE MAIN THERAPEUTIC AREAS IN THE DATABASE?

Biomarkers are not necessarily exclusive to one therapeutic area. They may be related to biological processes that are involved in many different diseases. When a biomarker from one of the core therapeutic areas is added to the database, information on all uses is added whether or not these fall within the main areas. Additionally, all biomarkers mentioned in sources published from 2009 onwards are added to the *Biomarkers Module*, whether or not they have uses in one of the main therapeutic areas.

INDICATION – SAFETY/TOX

WHAT IS THE STANDARD OF SELECTION TO INCLUDE AN ADVERSE EVENT IN THE BIOMARKERS MODULE?

A Safety/Tox use is created when there is a direct relationship demonstrated between an adverse event and the treatment received by the patients. For example, glutathione-S-transferase M1 as a biomarker of drug-induced liver toxicity or UGT1A1 as a biomarker for risk of irinotecan-induced neutropenia.

INDICATION – EXPERIMENTAL PATHOLOGY

WHAT IS THE STANDARD OF SELECTION TO INCLUDE A BIOMARKER USE IN THE EXPERIMENTAL PATHOLOGY CATEGORY?

An experimental pathology biomarker use is created for a biomarker that indicates the underlying cause of a condition or a toxic reaction. For example, VEGF as a biomarker of angiogenesis in cancer, HOXB1 as a biomarker of oxidative stress in Alzheimer's Disease or XRCC1 as a biomarker of the DNA damage response pathway in cancer.

WHAT ARE THE DEFINITIONS FOR ROLES IN THE BIOMARKERS MODULE?

Diagnosis: The role of this biomarker is to identify or detect a disease.

Disease Profiling: The role of this biomarker is to obtain information about the disease, but there is insufficient data to assign a clinical role. The data is often obtained from high throughput analyses, for example transcript profiling, and might be extrapolated to the processes that cause the disease.

Differential Diagnosis: The role of this biomarker is to distinguish between two or more diseases with similar signs and symptoms.

Monitoring Treatment Efficacy: The role of this biomarker is to identify signs of a change (usually beneficial) as a result of treatment. A biomarker used for monitoring treatment efficacy is usually measured before the treatment starts (baseline) and at stages throughout the treatment (follow up).

Monitoring Treatment Toxicity: The role of this biomarker is to identify signs of adverse effects as a result of treatment. Measured at baseline and at stages throughout treatment.

Prediction of Drug Resistance: The role of this biomarker is to detect possible resistance to a therapy and thus to exclude that therapy from the possible therapies available to the patient.

Predicting Treatment Efficacy: The role of this biomarker is to predict a probable beneficial outcome as a result of treatment.

Predicting Treatment Toxicity: The role of this biomarker is to predict a probable adverse effect as a result of treatment.

Prognosis: The role of this biomarker is to predict the probable outcome of a disease, i.e. how a patient's disease will progress and their chances of recovery. The prediction is based on the usual course of the disease seen in similar patients without therapy.

Monitoring Disease Progress: The role of this biomarker is to monitor the progress of a disease, usually for diseases for which there is no effective therapy.

Risk Stratification: The role of this biomarker is to determine a person's risk of suffering a particular clinical event within a specified period of time.

Risk Factor: The role of this biomarker is to determine a person's risk of a disease on the basis of epidemiological evidence.

Screening: The role of this biomarker is to sort a population into 'healthy' and 'non-healthy'. Screening is an epidemiological process, though the same process may serve for diagnosis as well.

Selection for Therapy: The role of this biomarker is to select a sub-group of patients suitable for a particular therapy.

Staging: The role of this biomarker is to describe how far a disease has progressed in a patient. The stage at diagnosis is often a prognostic indicator of overall survival and can be used as a guide for subsequent therapy.

Toxicity Profiling: The role of this biomarker is to obtain information about the underlying cause of an adverse or toxic event, but there is insufficient data to assign a predictive or monitoring role.

A biomarker use with "toxicity profiling" represents the birth of that use, often the first mention of the association between the biomarker and the adverse event and is always experimental. When new studies that focus on its predictive or monitoring role are added to our database, the role will be changed and upgraded.

HOW ARE VALIDITIES (LIFECYCLES PHASES) DEFINED IN THE BIOMARKERS MODULE?

Recommended/approved: Kit or measuring device/software is approved by the FDA, and/or the use of the biomarker is described in clinical practice guidelines or consensus development statements have been issued by clinical societies of international standing.

Late studies in humans: The biomarker use has been studied in humans (clinical trials and observational studies) but has not yet been approved by a regulatory authority. Should be studies with over 500 individuals. The objective is to measure efficacy and to establish the relationship found in early studies in humans.

Early Studies in Humans: The biomarker use has been studied in humans (clinical trials and observational studies). Should be studies with less than 500 individuals. The objective of these studies is to measure proof of concept and dose finding.

Experimental: Reports from preclinical studies (laboratory and/or animal studies).

Emerging: First mention of biomarker use, usually from patents and press releases

HOW ARE VALIDITIES (LIFECYCLES PHASES) ASSIGNED IN THE *BIOMARKERS MODULE*?

The highest validity corresponding to each use of a biomarker has to be assigned. This means if a use is reported in papers with results from clinical trials in different phases but the same use is described in a clinical practice guideline, the use will be assigned the validity status "Recommended/approved".

WHAT IS A PARAMETER IN THE *BIOMARKERS MODULE*?

A parameter defines which element of the biomarker was measured. The field is applied primarily to structural and physiological biomarkers: when the biomarker is structural, the parameter is geometric, e.g. volume, diameter, thickness; when the biomarker is physiological, the parameter is temporal, e.g. velocity, duration.

WHAT ABBREVIATIONS ARE USED FOR PARAMETERS IN THE *BIOMARKERS MODULE*?

- **FEF:** Forced Expiratory Flow
- **FEV:** Forced Expiratory Volume
- **FIV:** Forced Inspiratory Volume
- **PIF:** Peak Inspiratory Flow
- **SUV:** Standardized Uptake Value
- **TLC:** Total Lung Capacity

WHAT IS THE DIFFERENCE BETWEEN IMAGING TECHNOLOGIES AND IMAGING TECHNIQUES?

Technique terms are more specific than technology terms and are selected in preference to technology terms wherever possible. Technology terms refer to the machinery used to measure the biomarker (e.g. ultrasound (US)), whereas technique terms refer to the machinery and the manipulation of machinery and substrate (e.g. US-angiography).

WHAT ABBREVIATIONS ARE USED FOR IMAGING TECHNOLOGIES IN THE *BIOMARKERS MODULE*?

- **CT:** Computed Tomography
 - **EBCT:** Electron Beam Computed Tomography
 - **MDCT:** Multi-Detector Computed Tomography
 - **QCT:** Quantitative Computed Tomography
 - **sCT:** Spiral Computed Tomography

- **MR:** Magnetic Resonance
 - **EPRI:** Electron Paramagnetic Resonance Imaging
 - **MR-GILD:** Magnetic Resonance Gated Intracranial Liquor (CSF) Dynamics
 - **PEDRI:** Proton-Electron Double-Resonance Imaging
- **RN:** Radionuclide
 - **PET:** Positron Emission Tomography
 - **SPECT:** Single Photon Emission Computed Tomography
- **US:** Ultrasound
 - **EPI:** Echo Planar Imaging
- **X-ray**
 - **DPA:** Dual Photon Absorptiometry
 - **DXA:** Dual (energy) X-ray Absorptiometry
 - **SPA:** Single Photon Absorptiometry
 - **SXA:** Single X-ray Absorptiometry
 - **XRF:** X-ray Fluorescence

WHAT ABBREVIATIONS ARE USED FOR IMAGING TECHNIQUES IN THE *BIOMARKERS MODULE*?

- **DIR-MR:** Double Inversion Recovery- Magnetic Resonance
- **DTI:** Diffusion Tensor Imaging
- **DWI:** Diffusion Weighted Imaging
- **fMRI:** functional Magnetic Resonance Imaging
- **IVP:** Intravenous Pyelogram
- **IVUS:** Intravascular Ultrasound
- **MFC:** Magnetic Field Correlation
- **OCT:** Optical Coherence Tomography
- **PWI:** Perfusion Weighted Imaging
- **SWI:** Susceptibility Weighted Imaging
- **TDI:** Tissue Doppler Imaging

WHAT ARE THE CRITERIA FOR LINKING AN ORGANIZATION TO A *BIOMARKERS MODULE* RECORD?

Companies are linked to a biomarker record if they make a kit for which an FDA approval letter has been identified by our editorial analysts.

WHAT ARE THE CRITERIA FOR INCLUDING DIAGNOSTIC KITS IN THE *BIOMARKERS MODULE*?

Kits are included when there is an FDA 510(k) or an FDA PMA approval statement.

IS IT POSSIBLE TO SEARCH FOR A GENE VARIANT IN THE *BIOMARKERS MODULE*?

Genes are indexed in the *Biomarkers Module* using the name of the wild-type or predominant variant. Information on specific genomic variations such as SNPs is displayed in the corresponding use record.

There will also be more detail in the references related to a specific use of interest and by linking to related information in the Genomics Knowledge Area, for example in gene-related studies.

WHAT ARE THE DEFINITIONS FOR TYPES OF GENETIC VARIATIONS IN THE *BIOMARKERS MODULE*?

Polymorphism/mutation: describes a common change in the genetic code. These changes include single nucleotide polymorphisms; mutations (any change in the DNA of a cell); deletions (loss of DNA from a chromosome); and insertions (insertion of DNA into a chromosome). For example, the 405G>C polymorphism in VEGFA is a risk factor for endometriosis.

Gene duplication: describes a selective increase in the number of copies of a gene without a proportional increase in other genes. For example, duplication of the alpha-synuclein gene is a risk factor for Parkinson's Disease.

Epigenetic Change: describes changes in the regulation of the expression of gene activity without alteration of the DNA sequence. For example, the hypermethylation status of the bone morphogenetic protein 6 gene promoter is prognostic in patients with diffuse large B-cell lymphoma.

WHAT ARE THE CRITERIA FOR LINKING A "PRODUCT" TO A BIOMARKER USE RECORD?

Products are therapeutic or diagnostic agents. When a biomarker has been used to indicate a response to therapy, or depends on the administration of a diagnostic agent, the product is linked to the use.

WHAT TYPES OF "DIAGNOSTIC AGENTS" CAN I EXPECT TO SEE IN A BIOMARKER USE RECORD?

Diagnostic agents include contrast agents and stress agents used in physiological and imaging methods.

Contrast agent: a substance that induces a difference of image density between two areas (e.g. Pittsburgh Compound B).

Stress agent: a substance that disturbs the normal physiological equilibrium (e.g. adenosine). In combination with other diagnostic techniques, it can be used to determine a patient's response to stress.

HOW DOES A "PRODUCT MODIFIER" DIFFER FROM A "PRODUCT" IN THE *BIOMARKERS MODULE*?

A "Product Modifier" is linked to a biomarker when the product is specifically named in a clinical recommendation and/or an FDA 510(k) Approval statement. In contrast, a "Product" can be linked to a Use Record even if the use does not have the validity phase "recommended/approved".

WHAT ARE THE CRITERIA FOR LINKING A "MECHANISM OF ACTION" TO A BIOMARKER USE?

When a biomarker has been used to indicate a response to therapy with drugs categorized by their mechanism of action, then the mechanism of action is linked to the use.

HOW DOES A "MECHANISM MODIFIER" DIFFER FROM A "MECHANISM OF ACTION" IN THE *BIOMARKERS MODULE*?

A "Mechanism Modifier" is linked to a biomarker when the biomarker is a target for that mechanistic class of drugs. There is no such restriction for linking mechanisms of action directly to uses.

WHAT ARE THE CRITERIA FOR LINKING A "PRODUCT CATEGORY" TO A BIOMARKER USE?

A "Product Category" describes groupings into which bioactive compounds can be classified in terms of what they are, rather than how they work or what they are for. These include: chemical categories, biological factors, biotechnology medicines (antibodies, vaccines) and hormones.

When a biomarker has been used to indicate a response to therapy with bioactive compounds referred to by their "Product Category", the relevant product category term is linked to the use. For example, biomarker CA-125 is used for predicting the efficacy of treatment with platinum complexes in ovarian cancer patients. The product category "Platinum Complexes" will be linked to the relevant use record of biomarker CA-125.

WHAT ARE THE CRITERIA FOR LINKING A "THERAPEUTIC GROUP" TO A BIOMARKER USE?

A "Therapeutic Group" describes groups into which products can be classified in terms of what they are used for. These include broad groups such as "cardiovascular drugs", "neurological drugs"; or specific groups such as "antihypertensive drugs" or "antiarrhythmic drugs".

When a biomarker has been used to indicate a response to therapy with products described as belonging to a therapeutic group, the appropriate "therapeutic group" term will be linked to the use. For example, biomarker Cyclin D1 has been used to predict treatment efficacy of cancer immunotherapy agents in neuroblastoma patients. The therapeutic group "cancer immunotherapy" will be linked to the relevant use of biomarker Cyclin D1.

HEALTHCARE & SCIENCE REGIONAL OFFICES

North America

Philadelphia +1 800 336 4474
+1 215 386 0100
San Diego +1 858 273 8616

Latin America

Brazil +55 11 8370 9845
Other countries +1 215 823 5674

Europe, Middle East and Africa

Barcelona +34 93 459 2220
London +44 20 7433 4000

Asia Pacific

Singapore +65 6775 5088
Tokyo +81 3 5218 6500

For a complete office list visit:

science.thomsonreuters.com/contact

