Mitsui & Co. Global Strategic Studies Institute Deep Insight *August 2026*

HOW AI AND DATA ARE TRANSFORMING THE CLINICAL TRIAL ECOSYSTEM

— FROM A LABOR-INTENSIVE MODEL TO A DATA-DRIVEN MODEL —

Mika Kinoshita
Industrial Research Dept., Industrial & Social Studies Div.
Mitsui & Co. Global Strategic Studies Institute

SUMMARY

- In recent years, clinical trials have become increasingly complex as recruiting eligible patients has become more difficult and the number of endpoints has increased. As a result, labor-intensive trial operations are approaching their limits in terms of processing capacity and cost.
- Companies that provide solutions, such as AI- and healthcare data-enabled trial operations, patient recruitment, and data integration platforms, are gaining prominence as new key players in the clinical trial ecosystem.
- Pharmaceutical companies increasingly expect contract research organizations (CROs) to provide AI- and data-enabled solutions in addition to traditional labor-intensive services. The basis of competition is shifting toward the ability to simultaneously improve quality, speed, and productivity.

1. CHANGES IN THE ENVIRONMENT SURROUNDING CLINICAL TRIALS

1-1. Increasing trial complexity and development burden

The burden on pharmaceutical companies in new drug development is increasing year by year. The average time needed from the start of research to obtaining approval is 12 years, and development costs average $2.8 billion.[1] This growing burden reflects advances in precision medicine[2] and the expanding development of biopharmaceuticals. Because these development projects target narrowly defined patient populations, recruiting patients who meet the trial eligibility criteria is difficult. In fact, 10-30% of activated clinical trial sites reportedly fail to enroll a single patient.[3] Furthermore, the number of procedures and assessments per trial has tripled over the past 10 years to more than 250, and the number of data points collected has surged to approximately 3.5 million. In Phase 3 trials,[4] the final stage, the average number of endpoints has increased from 17 to 21 over the past 10 years.[5] This increasing complexity of clinical trials is affecting every stage of the process and is a contributing factor to the growing burden of drug development (Figure 1). Delays in clinical trials lead to delays in the launch of new drugs, and one estimate suggests that even a one-month delay results in more than $40

[1] CDER's annual novel drug approvals: 2013–2022, New Drug Therapy Approvals 2022, US Food & Drug Administration, current as of January 10, 2023.

[2] Precision medicine: A form of medicine that categorizes each patient in detail based on their genetic mutations and disease characteristics, and provides the most appropriate treatment for each individual.

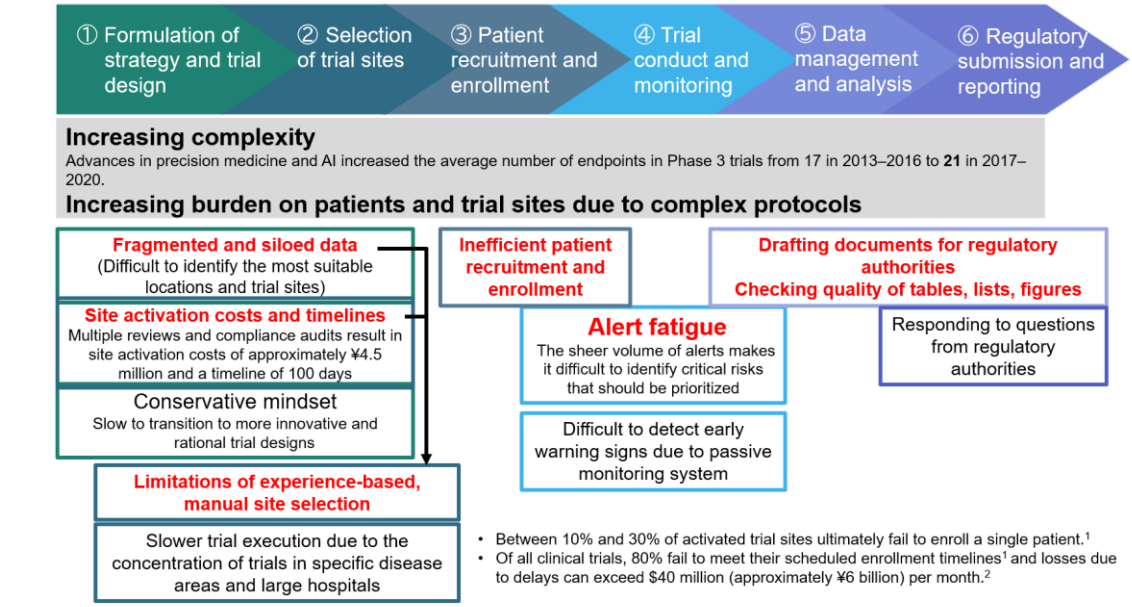
[3] McKinsey & Company, Unlocking peak operational performance in clinical development with artificial intelligence (January 2025)

[4] Phase 3 trial: A large-scale clinical trial conducted prior to submitting an application for a new drug to confirm the efficacy and safety of the drug in a large number of patients.

[5] McKinsey & Company, Accelerating clinical trials to improve biopharma R&D productivity (January 2024)

million in lost sales.[6] Therefore, improving the efficiency of clinical trials has become a critical business challenge for pharmaceutical companies.

Figure 1: Challenges posed by increasingly complex clinical trials



Source: Compiled by MGSSI based on data from McKinsey & Company, Unlocking peak operational performance in clinical development with artificial intelligence; Digital R&D the next frontier for biopharma[1]; Tufts Center for the Study of Drug Development, Tufts University, Quantifying the Value of a Day of Delay in Drug Development[2].

1-2. Limitations of the labor-intensive model

Traditionally, clinical trials have relied on labor-intensive operations centered on site visits by clinical research associates (CRAs)[7] and manual source data verification (SDV)[8]. However, as clinical trials have become more complex in recent years, the processing workload and the burden on clinical trial sites have both increased. It has also been reported that more than 66% of physicians who have served as clinical trial investigators subsequently choose not to participate in further trials, indicating that the labor-intensive model is reaching its limits.

1-3. Factors driving the expanding use of AI and data

Against the backdrop of the digitization of medical data, the widespread adoption of cloud infrastructure, and advances in generative AI and data analysis technologies based on large language models, the use of AI and data is rapidly expanding to address challenges in clinical trials. In addition, regulatory frameworks and pilot studies for the use of real-world data (RWD)[9] are advancing, particularly in Europe and the US, and researchers are exploring the application of RWD to areas such as trial design, patient identification, and external control arms (ECAs).[10] In response to these changes, pharmaceutical companies are accelerating partnerships with AI

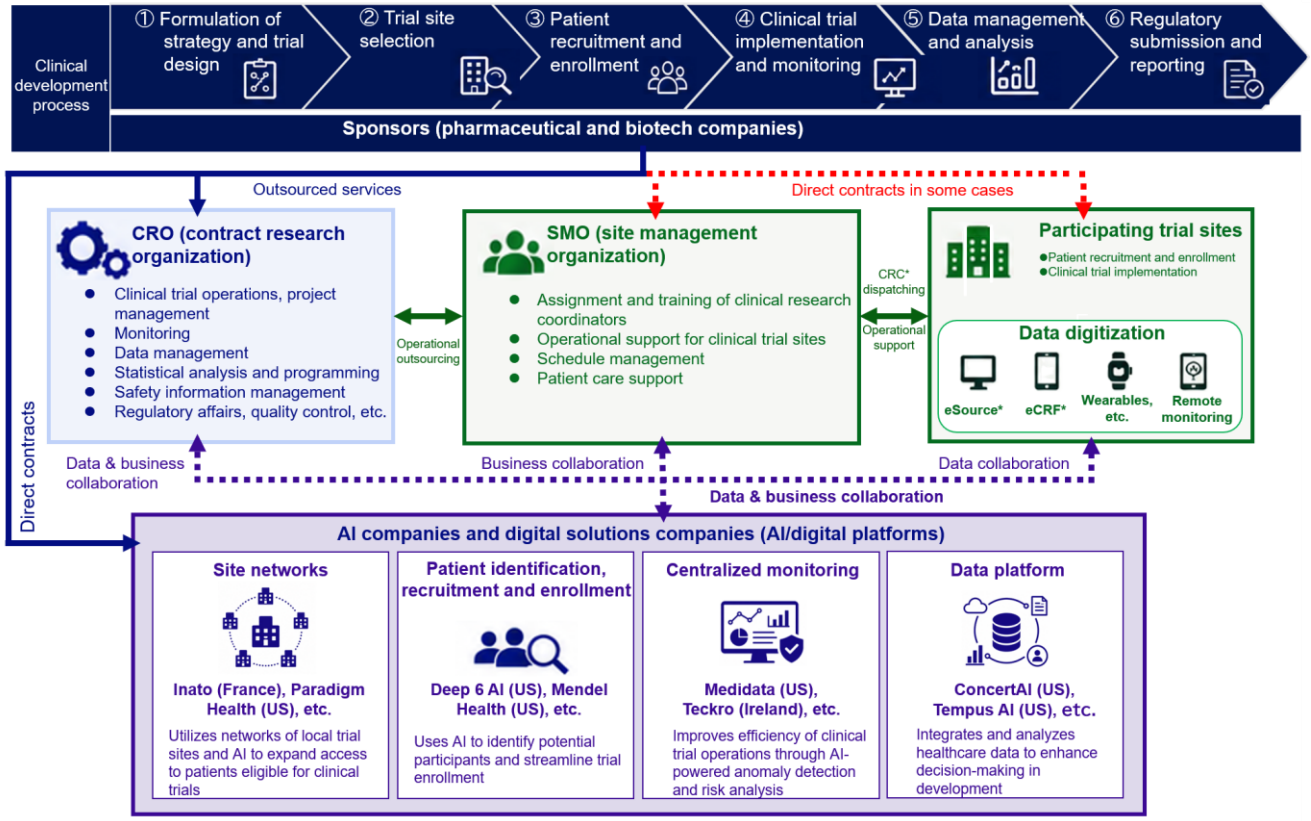
[6] Tufts Center for the Study of Drug Development, Tufts University, Quantifying the Value of a Day of Delay in Drug Development.
[7] Clinical research associate (CRA): A professional who visits clinical trial sites or monitors them remotely to ensure that clinical trials are being conducted appropriately in accordance with protocols and regulatory requirements, and who verifies data and provides support to those facilities.
[8] Source data verification (SDV): The process of verifying that source data, such as medical records and test results, match the data entered in the case report form.
[9] Real-world data (RWD): Medical data obtained from routine clinical practice, such as electronic medical records and insurance claims. The use of RWD in clinical trials and drug development is becoming increasingly widespread.
[10] External control arm (ECA): An approach that uses patient data obtained in real-world clinical settings or data from past clinical trials as a basis for comparison, which may reduce the number of patients that need to be newly enrolled as the control group. At present, ECAs do not replace control groups in all trials but are used as a complementary approach depending on the disease and trial design.

and data companies, in addition to investing in digital infrastructure, with the aim of streamlining development and improving success rates.

2. THE EVOLVING CLINICAL TRIAL ECOSYSTEM

In recent years, new solutions have emerged to address challenges at each stage of the clinical trial, and the structure of collaboration is changing. Besides major contract research organizations (CROs)[11] that have traditionally been involved in drug development, a diverse range of players, including AI companies, data companies, and companies operating networks of clinical trial sites (site networks),[12] are forming an ecosystem that supports clinical trials (Figure 2).

Figure 2: AI and digital infrastructure supporting the clinical trial ecosystem



*CRC: Clinical research coordinator
*eSource (electronic source data): Source data originally recorded and managed electronically, such as electronic health records and medical device data)
*eCRF (electronic case report form): Digital forms for the entry and management of clinical trial data

Source: Compiled by MGSSI based on various materials

2-1. Importance of patient recruitment and site networks

Patient recruitment is the biggest bottleneck affecting the duration of clinical trials and the timing of new drug launches, making rapid patient enrollment through site networks and site management organizations (SMOs)[13] crucial. Inato, a French site network company, has built a site network connecting more than 17,000 medical

[11] Contract Research Organization (CRO): A company that provides contracted services and support to pharmaceutical companies for drug development activities, such as clinical trials.
[12] Site network: A system that connects multiple hospitals and clinics into a network to facilitate coordinated patient referrals and clinical trials.
[13] Site Management Organization (SMO): A company that supports the conduct of clinical trials at medical institutions.

institutions in 80 countries. It links pharmaceutical companies and CROs with regional medical institutions, providing a platform that enables patients at these facilities to participate in clinical trials.[14,15] Similarly, Paradigm Health, a US site network company, has developed a system that integrates electronic medical records and genomic data from partner regional medical institutions with its own system to automatically identify eligible patients, thereby increasing the patient enrollment rate at participating community healthcare sites from approximately 4% to approximately 11%.[16] The US company Deep 6 AI specializes in patient recruitment, and has developed technology using natural language processing to analyze unstructured data, such as clinical notes in electronic medical records. Use of this technology at Cedars-Sinai Medical Center in the US has reduced the time required to identify eligible patients from six months to less than one hour.[17] These examples illustrate how expanded site networks, data integration, and AI are streamlining patient recruitment.

2-2. Centralized monitoring and automation of on-site operations

Monitoring activities, which involve verifying and managing the status of clinical trials, have become more sophisticated through the use of AI, enabling the automatic detection of anomalous data and improvements in the visualization and management of trial progress. As a result, these activities have evolved beyond simply verifying the status of trials to include early risk identification and management based on predictive analysis. Among AI companies specializing in clinical trial operations support, the US company Medidata Solutions is driving efficiency in trial operations by globally deploying a centralized monitoring platform that aggregates and visualizes data collected from individual trial sites in real time, enabling pharmaceutical companies and CROs to remotely monitor and analyze data across multiple sites.[14] The Irish company Teckro is providing on-site support using AI to digitize complex trial protocols, thereby significantly reducing investigators' cognitive burden and errors.[14] The US AI-driven data management company Saama Technologies provides a platform that uses AI assistants to automate clinical data cleaning and query management, thereby reducing manual tasks.[14]

2-3. Advances in data integration platforms and shifts in competitive dynamics

Data platforms that integrate electronic medical records, clinical trial data, real-world data (RWD), and other data sources are becoming a critical foundation for the design and management of clinical trials. ConcertAI, a US data platform company, uses AI to analyze integrated data in the field of oncology, thereby reducing the time required to develop clinical trial plans by 22–23% and supporting the development of ECAs. The company entered into a five-year strategic partnership with the US pharmaceutical company Bristol Myers Squibb, effective from 2022.[14,18] In addition, the US-based AI company Unlearn.ai generates "digital twins (virtual patients)" using machine learning and, since February 2022, has been partnering with Germany-based Merck to jointly develop technology that helps decrease the number of patients required for control groups.[14,18]

In response to the emergence of these data platforms, established global CROs are also working to strengthen their AI technologies and data infrastructure. Among leading global CROs, the US company Parexel is expanding its strategic partnership with US-based AI data analytics firm Palantir Technologies to build a clinical data analytics infrastructure leveraging an AI platform.[19] In addition, the US company Fortrea has established the AI Innovation Studio to develop AI solutions aimed at streamlining clinical trials, and is enhancing trial operations by integrating decentralized clinical trial (DCT)[20] platforms that support remote patient participation. Regionally focused CROs are also differentiating themselves by leveraging expertise in areas such as oncology and rare diseases, as well as regional networks, including those in Asia (Figure 3).

---

[14] Frost & Sullivan, Growth Opportunities and Innovative Use Cases for AI in Clinical Trials (December 2022).
[15] PitchBook Inato | Private Company Profile, Inato Website
[16] Paradigm Health Website (Solutions for Research Sites | Paradigm Health)
[17] Frost & Sullivan, Data Science Impacting the Pharmaceutical Industry Part II: Clinical Trial Applications (August 2020)
[18] Frost & Sullivan, Innovative AI-enabled Clinical Trial Companies: Strategic Profiling and Growth Opportunities (June 2024)
[19] Frost & Sullivan, Pharmaceutical Clinical Contract Peripheral Services, Global, 2025–2030 (March 2025).
[20] Decentralized Clinical Trial (DCT): Clinical trials in which patients do not visit medical institutions or other trial sites but instead participate from home or nearby healthcare facilities using online consultations, wearable devices, or other means.

Among AI companies, Tempus AI of the US is expanding its presence through extensive collaborations with pharmaceutical companies. Leveraging its strength in AI analytics platforms that utilize genomic and clinical data, in 2025, the company acquired the US company Deep 6 AI,[21] which has technology capable of efficiently identifying eligible patients by analyzing electronic medical records. Through this acquisition, Tempus AI is expanding its capabilities as a platform that supports activities ranging from the integration and analysis of healthcare data to patient recruitment.

Figure 3: Differentiation strategies of regionally focused CROs

Shift in value-added focus areas: Companies are differentiating themselves by focusing on vertical integration, specialized functions, and medical institution networks

Company	Strengths and areas of differentiation	Characteristics
Pharmaron (China)	CRDMO* model	Integrated services spanning drug discovery and preclinical research through manufacturing (CDMO*)
ClinChoice (US, China)	Regulatory affairs and submissions	Global regulatory and submission support across Japan, the US, Europe, and China
Prodia the CRO (Indonesia)	Integrated services	Indonesia's only CRO* with an integrated central laboratory and SMO* structure
Veeda (India)	Preclinical and bioanalytical services	Expanded capabilities in preclinical testing, bioanalysis, and central laboratory services
CMIC (Japan)	Network of trial sites + comprehensive support	Services extending to CDMO and sales support, supported by a network of thousands of medical institutions and extensive SMO and CRC* capabilities
EPS (Japan)	Network of trial sites + comprehensive support	

CRDMO (Contract Research, Development and Manufacturing Organization)
CDMO (Contract Development and Manufacturing Organization)
CRO (Contract Research Organization)
SMO (Site Management Organization)
CRC (Clinical Research Coordinator)
Source: Compiled by MGSSI

3. FUTURE BUSINESS OPPORTUNITIES AND IMPLICATIONS

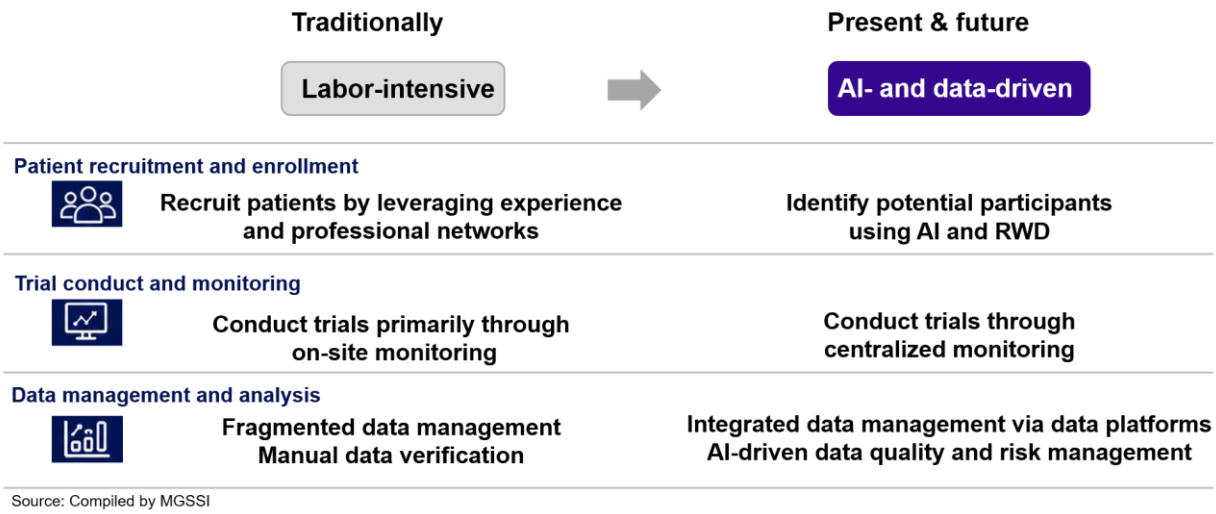
3-1. Shifts in value-added focus areas

The sources of value in clinical trials are shifting from labor-intensive operations that require the involvement of many specialists to advanced, specialized functions driven by AI and medical data (Figure 4). Approximately half (48%) of the clinical trials commissioned by major pharmaceutical companies involve function-specific outsourcing.19 In addition to full-service outsourcing to CROs, pharmaceutical companies are increasingly combining specialized services such as patient recruitment, data analysis, and regulatory support as needed.

Going forward, CRO competitiveness will depend not only on the ability to provide outsourced services but also on the ability to build on distinctive capabilities and combine them with advanced, AI- and data-enabled functions in areas such as trial operations, patient recruitment, and data analysis. The key will be whether CROs can help pharmaceutical companies improve the efficiency and quality of drug development.

[21] PitchBook, Deep 6 AI | Private Company Profile.

Figure 4: Shifting competitive factors in clinical trials



Source: Compiled by MGSSI

3-2. Implications for the Japanese market and business opportunities

In Japan, addressing “drug lag” and “drug loss”[22] has become a policy priority. Historically, when preparing to include Japan in multiregional clinical trials (MRCTs) initiated overseas, it was virtually standard practice to first conduct a domestic Phase 1 trial to confirm safety and appropriate dosage for Japanese patients. Because Phase 1 trials in Japan take a certain amount of time, there have been cases where Japan was unable to complete its Phase 1 trials in time for the start of Phase 3 trials, the final stage of global development, and was therefore excluded from the list of participating countries. This exclusion has been one of the factors contributing to delays in submitting applications for new drug approval (drug lag) due to the requirement for additional clinical trials in Japan and decisions not to pursue development in Japan (drug loss).

In light of these challenges, the regulatory approach was revised at the end of 2023, clarifying that, as a general rule, domestic Phase 1 trials are not required if existing data can demonstrate the safety and tolerability (a measure of the extent to which patients can tolerate a drug’s side effects) in Japanese patients. As a result, a regulatory environment is taking shape that makes it easier for pharmaceutical companies to include Japan in MRCTs at an early stage. To maximize the benefits of this regulatory reform, it is essential to build site networks capable of efficiently recruiting patients within Japan.

In Japan, while the use of electronic medical records and personal health records (PHRs)[23] is on the rise, data formats and input fields vary from one medical institution to another, and healthcare data are not yet readily usable across institutions. Therefore, to efficiently identify patients who meet clinical trial eligibility criteria, it is essential to promote the adoption of Fast Healthcare Interoperability Resources (FHIR),[24] an international standard for the interoperability of medical data across different healthcare IT systems, as well as to structure clinical data and establish mechanisms for secure data sharing among medical institutions. Building and providing a data infrastructure that standardizes and structures medical data currently scattered across various

[22] Drug lag and drug loss: “Drug lag” refers to the delay between the approval of a drug overseas and its approval and use in Japan, while “drug loss” refers to a situation where a drug, even though it has been approved and is in use overseas, is not developed or approved in Japan and therefore cannot be used there.
[23] Personal health record (PHR): Health and medical information managed by individuals. The information includes health screening results, medication history, and vital signs, and is expected to be utilized by medical institutions and in clinical trials.
[24] Fast Healthcare Interoperability Resources (FHIR): An international standard for enabling the interoperability of healthcare information. Developed by HL7 (Health Level Seven International), a US healthcare information standards organization

medical institutions, which would enable AI-enabled identification of patients who meet clinical trial criteria, is likely to present new business opportunities in Japan's clinical trial market.

Appendix: Key terminology

■ Clinical trial terms
- CDMO (Contract Development and Manufacturing Organization): Contract drug development and manufacturing organization
- CRA (Clinical Research Associate): Clinical trial monitor
- CRC (Clinical Research Coordinator): Clinical trial coordinator
- CRDMO (Contract Research, Development and Manufacturing Organization): Contract drug research, development, and manufacturing organization
- CRO (Contract Research Organization): Organization that provides outsourced clinical development services
- DCT (Decentralized Clinical Trial)
- eCRF (Electronic Case Report Form): Electronic form for capturing and managing clinical trial data
- eSource (Electronic Source Data): Source data originally recorded in electronic form, such as data from electronic health records and medical devices
- SDV (Source Data Verification): Verification of consistency between source data and case report form data
- SMO (Site Management Organization): Clinical trial site support organization

■ Data and AI terms
- FHIR (Fast Healthcare Interoperability Resources): International standard for healthcare data interoperability
- PHR (Personal Health Record): Health and medical information managed by individuals
- RWD (Real-World Data): Healthcare data generated through routine clinical practice, such as electronic medical records and claims data

Source: Compiled by MGSSI

Any use, reproduction, copying or redistribution of this report, in whole or in part, is prohibited without the prior consent of Mitsui & Co. Global Strategic Studies Institute (MGSSI). This report was created based on information and data obtained from sources believed to be reliable; however, MGSSI does not guarantee the accuracy, reliability, or completeness of such information or data. Opinions contained in this report represent those of the author and cannot in any way be considered as representing the unified opinion of MGSSI and Mitsui & Co. Group. MGSSI and Mitsui & Co. Group will not be liable for any damages or losses, whether direct or indirect, that may result from the use of this report. The information in this report is subject to change without prior notice.